26334 117 + Tpr Dudley

THE QUARTERLY JOURNAL OF
MATHEMATICS
OXFORD SERIES

Volume 3 *No.* 12 *December* 1932

CONTENTS

G. H. Hardy and J. E. Littlewood : Some Integral Inequalities connected with the Calculus of Variations 241

P. Dienes : Notes on Linear Equations in Infinite Matrices 253

E. Hopf : On certain Integral Equations. . . . 269

L. Carlitz : On a Problem in Additive Arithmetic (II) 273

A. G. Walker : The Second Curvature of a Sub-space 291

L. E. Lefevre : The Theory of the Oscillations of certain Stellar Models about their Steady-state Configurations 299

J. L. Burchnall : A Relation between Hypergeometric Series 318

OXFORD
AT THE CLARENDON PRESS
1932

Price 7s. 6d. net

THE QUARTERLY JOURNAL OF MATHEMATICS

OXFORD SERIES

Edited by T. W. CHAUNDY, W. L. FERRAR, E. G. C. POOLE
With the co-operation of A. L. DIXON, E. B. ELLIOTT, G. H. HARDY, A. E. H. LOVE, E. A. MILNE, E. C. TITCHMARSH

THE QUARTERLY JOURNAL OF MATHEMATICS (OXFORD SERIES) is, by arrangement with the publishers, the successor to *The Quarterly Journal of Mathematics* and *The Messenger of Mathematics*. The Journal is published in March, June, September, and December, at a price of 7s. 6d. net for a single number with an annual subscription (for four numbers) of 27s. 6d. post free.

Papers, of a length normally not exceeding 20 printed pages of the Journal, are invited on subjects of Pure and Applied Mathematics, and should be addressed 'The Editors, Quarterly Journal of Mathematics, Clarendon Press, Oxford'. Contributions can be accepted in French and German, if in typescript (formulae excepted). Authors of papers printed in the Quarterly Journal will be entitled to 50 free offprints. Correspondence on the *subject-matter* of the Quarterly Journal should be addressed, as above, to 'The Editors', at the Clarendon Press. All other correspondence should be addressed to the Publisher (Mr. Humphrey Milford, Oxford University Press, Amen House, Warwick Square, London, E.C. 4).

HUMPHREY MILFORD
OXFORD UNIVERSITY PRESS
AMEN HOUSE, LONDON, E.C. 4

SOME INTEGRAL INEQUALITIES CONNECTED WITH THE CALCULUS OF VARIATIONS

By G. H. HARDY AND J. E. LITTLEWOOD (*Cambridge*)

[Received 15 July 1932]

1. THE inequalities which we consider in this note may appear a rather miscellaneous collection; but they have this in common, that, unlike most of the inequalities of ordinary analysis, they depend upon, or are suggested by, problems in the calculus of variations. In this they resemble the inequality discussed by Bliss and by ourselves in vol. v of the *Journal of the London Mathematical Society*;† but the variational theory underlying this last inequality is decidedly more abstruse than any to which we shall appeal. The simpler problems discussed here should help to elucidate the difficult argument of Bliss.

Our proofs will be of three types, which we call A, B, C. Proofs of type A are strictly elementary; however the results may have been suggested, the proofs do not depend on variational theory. Proofs of type B depend on our knowledge of the form of an extremal, i.e. of a solution of the Euler equation associated with a variational problem, but are elementary in all other respects. It is natural, in view of the genesis of Euler's equation, that the principal weapon used in such proofs should be partial integration. Finally, proofs of class C depend essentially on the ideas and methods proper to the calculus of variations (as does the proof of Bliss to which we have referred). The interest of the paper lies largely in the logical relations between these different types of proof.

In order to avoid tiresome repetitions we state, once for all, that *whenever y and its derivative y'* (or f and f') *occur in a theorem or its proof, it is assumed that y is the integral of y'* (or f of f'). Similarly, if y'' occurs, y' is the integral of y''. This assumption is essential in all such variational problems.

We shall require the following lemma: *if y' is L^2 in $(0, a)$,‡ then $y = o(x^{\frac{1}{2}})$ for small x; and if $a = \infty$, then $y = o(x^{\frac{1}{2}})$ also for large x.*

† Bliss (**3**), Hardy and Littlewood (**6**).

‡ y' is measurable and its square integrable.

In the first place, by the inequality of Schwarz,

$$y^2 = \left\{\int_0^x y'(t)\,dt\right\}^2 \leqslant x\int_0^x y'^2\,dt = o(x)$$

for small x. Secondly, if $a = \infty$, we can choose $X = X(\epsilon)$ so that

$$\int_X^\infty y'^2\,dt < \epsilon^2;$$

and then

$$\{y(x)-y(X)\}^2 = \left\{\int_X^x y'(t)\,dt\right\}^2 \leqslant (x-X)\int_X^x y'^2\,dt < \epsilon^2 x,$$

$$|y(x)| < |y(X)|+\epsilon x^{\frac{1}{2}} < 2\epsilon x^{\frac{1}{2}},$$

for $x > x_0(\epsilon, X) = x_0(\epsilon)$.

2. We begin by proving

THEOREM 1. *If $\mu \geqslant 4$, $y(0) = 0$, $y(1) = 1$, and y' is L^2, then*

$$J = J(y) = \int_0^1 \left(\mu y'^2 - \frac{y^2}{x^2}\right) dy \geqslant \frac{2}{1-2a}, \tag{2.1}$$

where

$$a = \sqrt{\left(\frac{1}{4}-\frac{1}{\mu}\right)}.$$

There is equality if and only if $y = x^{\frac{1}{2}+a}$.

Our proof is of type *B*. We begin by supposing $\mu > 4$, so that $a > 0$. The Euler equation for J is

$$\mu x^2 y'' + y = 0, \tag{2.2}$$

and

$$y = \eta = x^{\frac{1}{2}+a}$$

is the extremal satisfying the conditions. There is equality in (2.1) when $y = \eta$.

We write generally

$$y = \eta + z$$

so that $z(0) = z(1) = 0$. Then

$$J(y) = J(\eta)+J(z)+2K(\eta, z),$$

where

$$K(\eta, z) = \int_0^1 \left(\mu\eta' z' - \frac{\eta z}{x^2}\right) dx.$$

Since z' is L^2, $z = o(x^{\frac{1}{2}})$ for small x. Hence, integrating by parts and observing that η satisfies (2.2), we find that $K = 0$ and

$$J(y) = J(\eta)+J(z). \tag{2.3}$$

It is therefore sufficient to prove that

$$J(z) > 0 \tag{2.4}$$

unless $z = 0$.

We now write $z = \eta u$,

so that $u(1) = 0$. Then

$$J_\delta(z) = \int_\delta^1 \left(\mu z'^2 - \frac{z^2}{x^2}\right) dx = \mu \int_\delta^1 (\eta u' + \eta' u)^2 \, dx - \int_\delta^1 \frac{\eta^2 u^2}{x^2} \, dx$$

$$= \mu \int_\delta^1 \eta^2 u'^2 \, dx + \int_\delta^1 \left(\mu \eta'^2 - \frac{\eta^2}{x^2}\right) u^2 \, dx + 2\mu \int_\delta^1 \eta\eta' uu' \, dx.$$

But

$$2\mu \int_\delta^1 \eta\eta' uu' \, dx = -\mu(\eta\eta' u^2)_\delta - \mu \int_\delta^1 (\eta'^2 + \eta\eta'') u^2 \, dx.$$

Combining these equations, and observing that

$$\eta' = (\tfrac{1}{2}+a)\frac{\eta}{x}, \qquad \mu\eta'' + \frac{\eta}{x^2} = 0, \qquad \eta u' = z' - \frac{\tfrac{1}{2}+a}{x} z,$$

and $z = o(x^{\frac{1}{2}})$, we obtain

$$J_\delta(z) = -\mu(\tfrac{1}{2}+a)\left(\frac{z^2}{x}\right)_\delta + \mu \int_\delta^1 (\eta u')^2 \, dx,$$

$$J(z) = \int_0^1 \left(\mu z'^2 - \frac{z^2}{x^2}\right) dx = \mu \int_0^1 \left(z' - \frac{\tfrac{1}{2}+a}{x} z\right)^2 dx. \tag{2.5}$$

This is positive unless the integrand is nul, when $z = C\eta$, which is impossible unless $C = 0$. This proves (2.4), and therefore the theorem, when $\mu > 4$.

It will be observed that (2.4) and (2.5) give

$$J(y) = \int_0^1 \left(\mu y'^2 - \frac{y^2}{x^2}\right) dx = \frac{2}{1-2a} + \mu \int_0^1 \left(y' - \frac{\tfrac{1}{2}+a}{x} y\right)^2 dx. \tag{2.6}$$

This identity may naturally be verified directly by partial integration when its form is observed. Since η has now disappeared, we may take $\mu = 4$, $a = 0$ in (2.5) and (2.6); and this completes the proof.

Trivial transformations of (2.6) give

$$\int_0^\xi \left(\mu y'^2 - \frac{y^2}{x^2}\right) dx = \frac{2}{1-2a} \frac{c^2}{\xi} + \mu \int_0^\xi \left(y' - \frac{\tfrac{1}{2}+a}{x} y\right)^2 dx, \tag{2.7}$$

where now $y(0) = 0$, $y(\xi) = c$. Suppose now that y' is L^2 in $(0, \infty)$, so that $c = y(\xi) = o(\xi^{\frac{1}{2}})$ for large ξ. Then, if we make $\xi \to \infty$, (2.7) becomes

$$\int_0^\infty \left(\mu y'^2 - \frac{y^2}{x^2}\right) dx = \mu \int_0^\infty \left(y' - \frac{\frac{1}{2}+a}{x} y\right)^2 dx.$$

Taking $\mu = 4$, we obtain

$$\int_0^\infty \left(4y'^2 - \frac{y^2}{x^2}\right) dx = 4 \int_0^\infty \left(y' - \frac{y}{2x}\right)^2 dx, \tag{2.8}$$

which is positive unless y is nul. This identity† gives the simplest proof of the inequality

$$\int_0^\infty \frac{y^2}{x^2} dx < 4 \int_0^\infty y'^2 dx. \tag{2.9}$$

3. Our proof of Theorem 1 is of type B; underlying it, there is a proof of type C. The family of extremals

$$y = \alpha\eta = \alpha x^{\frac{1}{2}+a}$$

defines a field‡ surrounding the extremal $y = \eta$. The slope-function of the field is

$$p = \alpha(\tfrac{1}{2}+a)x^{-\frac{1}{2}+a} = (\tfrac{1}{2}+a)\frac{y}{x}$$

and the Weierstrass excess-function is

$$\mathcal{E} = \mu(y'-p)^2.$$

Hilbert's invariant integral is

$$J^*(y) = \int \{(F - pF_p)\, dx + F_p\, dy\},$$

where

$$F = \mu p^2 - \frac{y^2}{x^2};$$

and the general theory tells us that

$$J(y) - J(\eta) = J(y) - J^*(\eta) = J(y) - J^*(y) = \int \mathcal{E}\, dx,$$

which is (2.6). It will be found that in this case

$$(F - pF_p)\, dx + F_p\, dy = d\left(\frac{y^2}{(\frac{1}{2}-a)x}\right).$$

† For similar, but less simple, identities for series, see Grandjot (**5**).

‡ See the treatises of Bolza (**4**) and Bliss (**2**). The field is 'uneigentlich', since all the curves pass through the origin.

4. Similar analysis may be applied to the more general inequality

$$\int_0^\infty \left(\frac{y}{x}\right)^k dx < \left(\frac{k}{k-1}\right)^k \int_0^\infty y'^k\, dx,$$

where $k > 1$, $y' \geqslant 0$.

THEOREM 2. *If*

$$k > 1, \qquad \frac{1}{k}+\frac{1}{k'} = 1, \qquad \mu \geqslant \left(\frac{k}{k-1}\right)^k = K,$$

$y(0) = 0$, $y(1) = 1$, *and* y' *is positive and* L^k, *then*

$$J(y) = \int_0^1 \left(\mu y'^k - \frac{y^k}{x^k}\right) dx \geqslant \frac{1}{(k-1)(1-\lambda)},$$

where λ *is the* (unique) *number defined by*

$$\mu(k-1)\lambda^{k-1}(\lambda-1)+1 = 0, \qquad \lambda > 1/k'.$$

The underlying identity, corresponding to (2.6), is

$$\int_0^1 \left(\mu y'^k - \frac{y^k}{x^k}\right) dx = \frac{1}{(k-1)(1-\lambda)} + \\ + \mu \int_0^1 \left\{y'^k - \left(\frac{\lambda y}{x}\right)^k - k\left(y' - \frac{\lambda y}{x}\right)\left(\frac{\lambda y}{x}\right)^{k-1}\right\} dx;$$

and the integral on the right-hand side is positive, unless $y' \equiv \lambda y/x$. When $\mu = K$, $\lambda = 1/k'$ and $\mu\lambda^k = 1$.

5. THEOREM 3. *If* $y(0) = y(1) = 0$ *and* y' *is* L^2, *then*

$$\int_0^1 \frac{y^2}{x(1-x)}\, dx \leqslant \tfrac{1}{2} \int_0^1 y'^2\, dx,$$

with equality only when $y = Cx(1-x)$.

Here
$$J(y) = \int_0^1 \left(\tfrac{1}{2}y'^2 - \frac{y^2}{x(1-x)}\right) dx$$

and Euler's equation is

$$y'' + \frac{2y}{x(1-x)} = 0.$$

The general solution is

$$y = x(1-x)\left(\alpha + \beta \log \frac{x}{1-x}\right),$$

and the family $y = \alpha x(1-x)$

defines a field† of extremals round any one of its members.

Here

$$p = \alpha(1-2x) = \frac{1-2x}{x(1-x)}\,y,$$

$$(F-pF_p)\,dx + F_p\,dy = d\left(\frac{1-2x}{2x(1-x)}\,y^2\right),$$

and the identity to which we are led is

$$\int_0^1 \left(\tfrac{1}{2}y'^2 - \frac{y^2}{x(1-x)}\right) dx = \frac{1}{2}\int_0^1 \left(y' - \frac{1-2x}{x(1-x)}\right)^2 dx.$$

Our proof is of type C or B, according as we construct this identity from the general theory or verify it directly.

Theorem 3 was suggested to us by the following theorem of Pólya.‡

THEOREM 4. *If $x(1-x)y'^2$ is integrable, then*

$$0 \leqslant \int_0^1 y^2\,dx - \left(\int_0^1 y\,dx\right)^2 \leqslant \tfrac{1}{2}\int_0^1 x(1-x)y'^2\,dx.$$

The sign of equality can occur in the two possible places only when y is (i) *constant*, (ii) *linear*.

Pólya's proof is as follows. The first inequality follows at once from Schwarz's inequality. For the second, we have

$$\begin{aligned}\int_0^1 y^2\,dx - \left(\int_0^1 y\,dx\right)^2 &= \tfrac{1}{2}\int_0^1\int_0^1 (y(u)-y(v))^2\,du\,dv \\ &= \int_0^1 du \int_u^1 dv \left(\int_u^v y'(t)\,dt\right)^2 \leqslant \int_0^1 du \int_u^1 (v-u)\,dv \int_u^v y'^2\,dt \\ &= \int_0^1 y'^2\,dt \int_0^t du \int_t^1 (v-u)\,dv = \tfrac{1}{2}\int_0^1 t(1-t)y'^2\,dt,\end{aligned}$$

with equality only when y' is constant.

Theorems 3 and 4 can also be deduced from 'Parseval's theorem' in the theory of Legendre series. A trivial transformation reduces them to the forms

$$\int_{-1}^1 \frac{y^2}{1-x^2}\,dx \leqslant \tfrac{1}{2}\int_{-1}^1 y'^2\,dx, \tag{5.1}$$

† Here 'uneigentlich' at both ends. ‡ Communicated to us by letter.

with $y(-1) = y(1) = 0$; and

$$\int_{-1}^{1} y^2\,dx \leqslant \tfrac{1}{2}\int_{-1}^{1} (1-x^2)y'^2\,dx, \tag{5.2}$$

with

$$\int_{-1}^{1} y\,dx = 0.$$

To prove (5.1), we suppose that

$$y' \sim \sum a_n P_n(x),$$

with $a_0 = 0$. Then

$$y = \sum a_n \int_{-1}^{x} P_n(t)\,dt = \sum a_n p_n(x),$$

say. It can easily be verified that

$$\int_{-1}^{1} \frac{p_m p_n}{1-x^2}\,dx = 0 \quad (m \neq n), \qquad = \frac{2}{n(n+1)(2n+1)} \quad (m = n),$$

and (5.1) then follows from

$$\sum_{1}^{\infty} \frac{2}{n(n+1)} \frac{a_n^2}{2n+1} \leqslant \sum_{1}^{\infty} \frac{a_n^2}{2n+1}.$$

Similarly, if we suppose that

$$y \sim \sum a_n P_n(x),$$

and use the formulae

$$\int_{-1}^{1} (1-x^2)P_m' P_n'\,dx = 0 \quad (m \neq n), \qquad = \frac{2n(n+1)}{2n+1} \quad (m = n),$$

(5.2) will be found to follow from

$$\sum_{1}^{\infty} \frac{a_n^2}{2n+1} \leqslant \sum_{1}^{\infty} \frac{n(n+1)}{2} \frac{a_n^2}{2n+1}.$$

In each case there is equality only when $a_2 = a_3 = \ldots = 0$.

6. THEOREM 5. *If $y(0) = 0$, $y(1) = 1$, k is a positive integer, and y' is L^{2k}, then*

$$\int_{0}^{1} y^{2k}\,dx \leqslant C \int_{0}^{1} y'^{2k}\,dx,$$

where

$$C = \frac{1}{2k-1}\left(\frac{\pi}{2k}\operatorname{cosec}\frac{\pi}{2k}\right)^{-2k}.$$

There is equality only for a certain hyperelliptic y.

If $a = 1/C$ and

$$J(y) = \int_0^1 (y'^{2k} - ay^{2k})\, dx,$$

Euler's equation is

$$(2k-1)y'^{2k-2}y'' + ay^{2k-1} = 0,$$

which gives

$$(2k-1)y'^{2k} = b - ay^{2k}.$$

If we take $b = a$, we obtain the special family of extremals

$$x = \left(\frac{2k-1}{a}\right)^{1/2k} \int \frac{dy}{(1-y^{2k})^{1/2k}},$$

and then, since

$$\int_0^1 \frac{dy}{(1-y^{2k})^{1/2k}} = \frac{1}{2k} \int_0^1 (1-u)^{-1/2k} u^{1/2k-1}\, du = \frac{\pi}{2k} \operatorname{cosec} \frac{\pi}{2k},$$

the extremal through the origin will also pass through $(1, 1)$.

If we denote this extremal by $y = \eta$, and observe that $\eta'(1) = 0$, we obtain

$$J(\eta) = [\eta\eta'^{2k-1}]_0^1 - \int_0^1 \eta\{(2k-1)\eta'^{2k-2}\eta'' + a\eta^{2k-1}\}\, dx = 0,$$

and it is plain that the theorem will follow if $y = \eta$ gives a strong minimum for $J(y)$. That this is so follows from the general theory. The graph of η is a rising curve whose shape resembles generally that of the curve $y = \sin \frac{1}{2}\pi x$, to which it reduces when $k = 1$. The family

$$y = \alpha\eta$$

defines a field, and the Weierstrass function

$$y'^{2k} - p^{2k} - 2k(y' - p)p^{2k-1}$$

is positive; and this is what is required for the application of the general theory.

This proof is of type C and (in view of the trouble of calculating the slope-function p in terms of x and y) it might be difficult to construct a much more elementary proof. When $k = 1$, the inequality reduces to

$$\int_0^1 y^2\, dx \leqslant \frac{4}{\pi^2} \int_0^1 y'^2\, dx,$$

with equality for $y = \sin \frac{1}{2}\pi x$. This inequality is very like one due to Wirtinger, which has an important application to the proof of the isoperimetrical property of the circle.† In this case there are more elementary proofs, and the reader will find it instructive to construct proofs similar to those of the previous sections, and to deduce the result from the theory of Fourier series.

7. In § 6 we used a proof of type C which it might be difficult to reduce to type B. In this section we use a proof of type B which we cannot replace by one of type A.

THEOREM 6. *If y and y'' are L^2, then*

$$\left(\int_0^\infty y'^2\,dx\right)^2 \leqslant 4\int_0^\infty y^2\,dx \int_0^\infty y''^2\,dx, \tag{7.1}$$

with equality only when

$$y = \eta = Ce^{-x\cos\gamma}\sin(x\sin\gamma - \gamma), \qquad (\gamma = \tfrac{1}{3}\pi). \tag{7.2}$$

The Euler equation for the 'isoperimetrical' problem of making

$$J_1 = \int_0^\infty y'^2\,dx$$

a maximum, when

$$J_0 = \int_0^\infty y^2\,dx, \qquad J_2 = \int_0^\infty y''^2\,dx$$

are fixed, is of the form

$$ay'''' + by'' + cy = 0,$$

and its solutions are linear combinations of real or complex exponentials. When we tried to choose the parameters of the solutions in the most advantageous way, we were led to select the function (7.2).

We shall deduce Theorem 6 from

THEOREM 7. *If the conditions of Theorem 6 are satisfied, then y' is L^2 and*

$$J(y) = \int_0^\infty (y^2 + y''^2 - y'^2)\,dx \geqslant 0,$$

with equality only for $y = \eta$.

† See Blaschke (**1**), 105.

In the first place

$$\int_0^X y'^2\,dx = [yy']_0^X - \int_0^X yy''\,dx. \tag{7.3}$$

Since J_0 and J_2 are finite, the last integral tends to a finite limit when $X \to \infty$. Hence, if J_1 were infinite, yy', and *a fortiori* $\frac{1}{2}y^2 = \int yy'\,dx$, would tend to $+\infty$ with x, which is inconsistent with the convergence of J_0. Thus J_1 is finite, and all three terms in (7.3) tend to limits. In particular yy' tends to a limit, which can only be zero.

It is easy to verify that

$$\eta+\eta'+\eta'' = 0, \qquad \eta(0)+\eta'(0) = -\eta''(0) = 0,$$

and

$$J(\eta) = 0.$$

We now write

$$y = c\eta+z,$$

choosing c so that $z(0) = 0$. Since z differs from y by a function which, with its derivatives, vanishes exponentially at infinity,

$$zz' \to 0, \qquad (\eta+\eta')z \to 0, \qquad (\eta+\eta')z' \to 0.$$

Also

$$J(y) = c^2J(\eta)+2cK(\eta,z)+J(z),$$

and

$$\begin{aligned} K &= \int_0^\infty (\eta z+\eta''z''-\eta'z')\,dx \\ &= -\int_0^\infty (\eta'+\eta'')z\,dx - \int_0^\infty (\eta+\eta')z''\,dx - \int_0^\infty \eta'z'\,dx \\ &= \int_0^\infty (\eta+\eta'+\eta'')z'\,dx = 0. \end{aligned}$$

Hence (7.4) reduces to $J(y) = J(z)$, and it is sufficient to prove that

$$J(z) > 0$$

unless $z = 0$. But, since $z(0) = 0$ and $zz' \to 0$,

$$\int_0^\infty z'^2\,dx = -\int_0^\infty zz''\,dx,$$

and so

$$J(z) = \int_0^\infty (z^2+zz''+z''^2)\,dx > 0.$$

This proves Theorem 7. To deduce Theorem 6, we replace x by λx and $y(\lambda x)$ by $y(x)$, when we obtain

$$\lambda^4 J_0 - \lambda^2 J_1 + J_2 = \int_0^\infty (\lambda^4 y^2 - \lambda^2 y'^2 + y''^2)\, dx > 0.$$

Since this is true for all λ, $J_1^2 < 4J_0 J_2$.

Theorem 6 suggests that, if $k > 1$,

$$\left(\int_0^\infty |y'|^k\, dx\right)^2 \leqslant K(k) \int_0^\infty |y|^k\, dx \int_0^\infty |y''|^k\, dx.$$

We cannot deal with this inequality completely; it is easy to prove $K(k)$ finite for every k, but we cannot determine its best possible value, even for $k = 4$. A theorem of Landau† suggests that the limit of $K^{1/k}$ when $k \to \infty$ should be 4.

8. It may be interesting to remark in conclusion that the solution of the corresponding problem for the interval $(-\infty, \infty)$ is simpler and quite different.

THEOREM 8. *If y and y'' are L^2 in $(-\infty, \infty)$, then*

$$\left(\int_{-\infty}^\infty y'^2\, dx\right)^2 < \left(\int_{-\infty}^\infty y^2\, dx\right)\left(\int_{-\infty}^\infty y''^2\, dx\right)$$

unless y is nul. The constant 1 is the best possible.

For, as in § 7, $yy' \to 0$ when x tends to infinity, and

$$\int_{-\infty}^\infty y'^2\, dx = -\int_{-\infty}^\infty yy''\, dx,$$

so that the result follows from Schwarz's inequality.

To prove the constant the best possible, take first

$$y = \sin x \quad (|x| \leqslant n\pi), \qquad = 0 \quad (|x| > n\pi),$$

and then round off the corners at $x = \pm n\pi$ so as to make y'' continuous. We can do this with changes in J_0, J_1, and J_2 which are independent of n. Then all three integrals differ from $n\pi$ by numbers independent of n, and

$$J_1^2 > (1-\epsilon) J_0 J_2$$

for every ϵ and sufficiently large n.

† Landau (7).

REFERENCES

1. W. Blaschke, *Kreis und Kugel* (Leipzig, 1916).
2. G. A. Bliss, *Calculus of Variations* (Chicago, 1924).
3. —— 'An integral inequality', *Journal London Math. Soc.*, 5 (1930), 40–6.
4. O. Bolza, *Vorlesungen über Variationsrechnung* (Leipzig, 1909).
5. K. Grandjot, 'On some identities relating to Hardy's convergence theorem', *Journal London Math. Soc.*, 3 (1928), 114–17.
6. G. H. Hardy and J. E. Littlewood, 'Notes on the theory of series (XII): On certain inequalities connected with the calculus of variations', *Journal London Math. Soc.*, 5 (1930), 34–9.
7. E. Landau, 'Einige Ungleichungen für zweimal differentiierbare Funktionen', *Proc. London Math. Soc.* (2), 13 (1914), 43–9.

NOTES ON LINEAR EQUATIONS IN INFINITE MATRICES

By P. DIENES (*London*)

[Received 16 December 1931]

1. The problem. Definitions. For given y_n the solution of

$$\sum_{k=1}^{\infty} a_{nk} x_k = y_n \qquad (n = 1, 2, \ldots, \infty) \tag{1}$$

hinges on certain matrix characteristics of the table a_{nk}. The same equations are used, in the theory of infinite quadratic forms, for the transformation of the variables x_k into y_n in order to reduce $\sum_{n,k=1}^{\infty} a_{nk} x_n x_k$ to a diagonal form $\sum_{i=1}^{\infty} b_i y_i^2$.

Again, the generalized A-limit of x_k is that of y_n in (1) whenever the latter exists. In this connexion, $A = (a_{nk})$ being given, the problem of the consistency of the various generalizations of limit leads to the problem of finding a matrix Ω and a diagonal matrix D such that

$$A = \Omega D \Omega^{-1}. \tag{2}$$

In the Heisenberg-Dirac theory of atom mechanics the central problem is to find an Ω and a D such that

$$\Omega^{-1} A \Omega = D. \tag{3}$$

The determination of Ω in (2) and (3) requires, as a rule, the solution of

$$AX - XD = 0. \tag{4}$$

Also, the fundamental equation for quantization is

$$AX - XA = 1. \tag{5}$$

Therefore we thought it desirable to collect the more or less obvious results on linear matrix equations, with illustrative examples, as an introduction to the theories mentioned above. Some results, especially in §§ 4, 5, seem to be new. A list of papers on the subject is appended for reference. The point of view adopted, first put forward in A. C. Dixon† (1902) for certain systems (1), aims at establishing facts about the existence or non-existence of solutions and at developing methods for explicit formulae *without making use of*

† References are to the bibliography at the end of the paper.

infinite determinants. For functions of a matrix variable see D. Hilbert (1909), Riesz, chap. 5, and Dienes (1930). I owe some valuable suggestions to Professor A. C. Dixon.

DEFINITIONS. An infinite matrix is a twofold table

$$A = (a_{nk}) \qquad (n, k = 1, 2, \ldots, \infty)$$

of real or complex numbers, with addition and multiplication defined by

$$A+B = (a_{nk}+b_{nk}), \qquad cA = (ca_{nk}), \qquad AB = \Big(\sum_{i=1}^{\infty} a_{ni}b_{ik}\Big), \quad (6)$$

where c is any real or complex number (scalar). The first suffix n specifies the row, the second suffix the column. The zero matrix 0 is the table $a_{nk} = 0$. The elements δ_k^n of the unit matrix, denoted by 1 or E, are

$$\delta_k^n = \begin{cases} 0 \text{ if } n \neq k \\ 1 \text{ if } n = k. \end{cases} \quad (7)$$

These two definitions are justified by the obvious equations

$$0A = A0 = 0, \qquad EA = AE = A.$$

We also notice that the matrices cE form a sub-algebra which is simply isomorphic with complex algebra, so that matrix algebra can be considered as an extension of complex algebra.

The elements a_{nn} form the (principal) diagonal. If, outside the diagonal, $a_{nk} = 0$, then A is a *diagonal* matrix denoted by $D, \Delta, \Theta, \Gamma$. If $a_{nk} = 0$ above (below) the diagonal, then A is a *lower* (*upper*) *semi-matrix*. If every row (column) contains a finite number of non-zero elements as, for example, in lower (upper) semi-matrices, the matrix is *row-finite* (*column-finite*). If $a_{nk} = 0$ outside a certain finite number of rows (columns), then A is *row-bounded* (*column-bounded*). For example, (a_{n1}) denotes the matrix whose first column is a_{n1}, all the other elements being 0.

If $AB = 1$, then B is a *rh.* (right-hand) *reciprocal* of A denoted by A^{-1}, and A is a *lh.* (left-hand) *reciprocal* of B, denoted by ${}^{-1}B$. Similarly, if both A and B are different from 0 and if $AB = 0$, then B is a *rh.* 0-*divisor* of A, denoted by A^0, and A is a *lh.* 0-*divisor* of B, denoted by ${}^0B$.

2. Remarks on matrix algebra. The sum of two matrices always exists, and

$$A+B = B+A, \qquad (A+B)+C = A+(B+C),$$

but products are in general not commutative and AB may not exist

even when BA exists. For example, if $B = (b_{n1})$, then $BA = (b_{n1}a_{1k})$ exists for arbitrary b_{n1} and A; on the other hand

$$(AB)_{n1} = \sum_{i=1}^{\infty} a_{ni}b_{i1}, \qquad (AB)_{nk} = 0, \text{ if } k > 1,$$

so that AB does not exist when $\sum_i a_{ni}b_{i1}$ is divergent.

The distributive law

$$A(B+C) = AB+AC, \qquad (B+C)A = BA+CA$$

holds in the sense that, if AB and AC exist, then also $A(B+C)$ exists and is equal to $AB+AC$. But $A(B+C)$ may exist when AB and AC do not exist. For example, if

$$a_{nk} = 1 \text{ for every } n \text{ and } k,$$

$$b_{nk} = n^{-2}+1,\ c_{nk} = n^{-2}-1 \text{ for every } k.$$

The product law may not hold even for simple matrices. For example, if in D and Δ, $d_{2n} = 0$, $\delta_{2n+1} = 0$ and the other diagonal elements are different from 0, then we have $D\Delta = (d_n\delta_n) = 0$, though neither D nor Δ is the zero matrix.

Multiplication is not necessarily associative. For example, if $a_{nk} = 1$, $c_{nk} = 1$ for every n and k, and if

$$\sum_{j=1}^{\infty}\Big(\sum_{i=1}^{\infty} b_{ij}\Big) \neq \sum_{i=1}^{\infty}\Big(\sum_{j=1}^{\infty} b_{ij}\Big),$$

then
$$(AB)C = \sum_{j=1}^{\infty}\Big(\sum_{i=1}^{\infty} b_{ij}\Big) \neq A(BC) = \sum_{i=1}^{\infty}\Big(\sum_{j=1}^{\infty} b_{ij}\Big).$$

As a numerical example we put

$$b_{ij} = \frac{i^2}{i^2+j} - \frac{i^2}{i^2+j-1} - \frac{(i-1)^2}{(i-1)^2+j} + \frac{(i-1)^2}{(i-1)^2+j-1},$$

in which case the partial sums are

$$S_{ij} = \frac{i^2}{i^2+j},$$

so that the two different summations lead to two different sums.

We shall say that a product is associative if (i) the product exists for every succession of the multiplications involved, the order of the factors remaining the same, (ii) all the products so obtained are equal. If, in a class of matrices, every finite product is associative, we say that the class is an associative field F. Diagonal matrices form an associative (and commutative) field.

Bound of a matrix. If, by a given rule, a number $|A| \geqslant 0$ is assigned to every matrix of a given class satisfying the conditions

(i) $|cA| = |c||A|$, $|E| = 1$ (c any complex number),

(ii) $|A+B| \leqslant |A|+|B|$,

(iii) $|AB| \leqslant |A||B|$,

we say that the class is *bounded* according to the given rule, and $|A|$ is the corresponding *bound* or *modulus* of A. For a similar definition see Helly, p. 61. Such a bound is the generalization of absolute value (modulus) of complex numbers, which justifies the notation. We notice that $|a_{nk}|$ denotes the absolute value of the complex number a_{nk}, so that, in general, $|a_{nk}| \neq |A|$.

The bound is said to be *regular* if it also satisfies the conditions

(iv) $|a_{nk}| \leqslant |A|$,

(v) $|D| = \bar{d}$,

where $\bar{d}$ is the finite upper bound of $|d_n|$. Condition (v) means that, if the upper bound $\bar{d}$ of $|d_n|$ is finite, every regular bound of the corresponding diagonal matrix D is exactly $\bar{d}$.

We notice that

$$\left|\sum_{p=1}^{\infty} c_p A_p\right| \leqslant \sum_{p=1}^{\infty} |c_p||A_p|, \tag{1}$$

whenever the right-hand side exists, and in this case, if A_1, A_2,... belong to the same associative field F, then $\sum c_p A_p$ also belongs to F.

Examples. Putting $\sum_{k=1}^{\infty} |a_{nk}| = M_n$, $\sum_{n=1}^{\infty} |a_{nk}| = N_k$, the upper bound of $M_n(N_k)$, when it is finite, is the $K_r(K_c)$ *bound* of A (see Dienes (1931), chap. 12).

The finite upper bound of $\left|\sum_{n,k=1}^{\infty} a_{nk} x_n x_k\right|$, for sequences x_n such that $\sum_{n=1}^{\infty} |x_n|^2 = 1$, is the H (Hilbert) *bound* of A. It follows that

$$\sum_{n=1}^{\infty} |a_{nk}|^2 \leqslant |A|^2, \qquad \sum_{k=1}^{\infty} |a_{nk}|^2 \leqslant |A|^2,$$

but the converse is not necessarily true.

It is readily verified that the K_r, K_c, and H bounds are regular.

We finally notice that row-finite, column-finite, K_r, K_c, and H matrices, i.e. the class of matrices having a finite K_r, K_c, or H bound respectively, form so many associative fields.

In fact, the product of two row-finite (column-finite) matrices is a row-finite (column-finite) matrix, and, in

$$\sum_{j=1}^{\infty}\Big(\sum_{i=1}^{\infty}a_{ni}b_{ij}\Big)c_{jk}=\sum_{i=1}^{\infty}a_{ni}\Big(\sum_{j=1}^{\infty}b_{ij}c_{jk}\Big), \tag{2}$$

the infinite sums reduce to finite sums.

If A, B, C are K_r matrices, say,

$$\sum_{i=1}^{\infty}|a_{ni}b_{ik}|\leqslant|B|\sum_{i=1}^{\infty}|a_{ni}|\leqslant|A||B|$$

and

$$\sum_{j=1}^{\infty}\Big(\sum_{i=1}^{\infty}|a_{ni}b_{ij}|\Big)|c_{jk}|\leqslant|C|\sum_{j=1}^{\infty}\sum_{i=1}^{\infty}|a_{ni}b_{ij}|\leqslant|C||B|\sum_{i=1}^{\infty}|a_{ni}|\leqslant|A||B||C|,$$

so that (2) is absolutely convergent.

For H matrices see Hilbert (1912), p. 129.

3. Reciprocals of matrices. The second of the equations

$$AX=B,\qquad XA=B \tag{1}$$

may be written in the form $A^*X^*=B^*$, where the star indicates the transposed matrix (rows and columns exchanged); we therefore deal with the first equation only.

If a diagonal element in D is zero, $DX=1$, i.e. $d_n x_{nk}=\delta_k^n$ has no solution, since, if $d_{n'}=0$, say, $d_{n'}x_{n'n'}=1$ has none. Thus (1) may have no solution. However, if A^{-1} is a solution of

$$AX=1, \tag{2}$$

$A^{-1}B$ is that of (1) provided $AA^{-1}B$ is associative. Therefore we shall first examine (2).

If A^{-1} is a solution of (2), $A^{-1}+A^0$ is another solution, and every solution is of that form. Moreover, in the same associative field, no A has an A^{-1} and an ${}^0A$, or an ${}^{-1}A$ and an A^0, since $({}^0AA)A^{-1}=0$ and ${}^0A(AA^{-1})={}^0A$.

The relation between lh. and rh. reciprocals is illustrated by I. (i) *If A^{-1} is unique and $AA^{-1}.A$ is associative, then A^{-1} is also a lh. reciprocal of A and the only one for which ${}^{-1}AAA^{-1}$ is associative.* (ii) *If A has both an A^{-1} and an ${}^{-1}A$, and ${}^{-1}AAA^{-1}$ is associative, then ${}^{-1}A=A^{-1}$, and A has no other two-sided reciprocal for which ${}^{-1}AAA^{-1}$ is associative.*

PROOF. (i) If Ω is a solution of $AX=B$, all the other solutions are of the form $\Omega+A^0$. In particular, the solutions of $AX=A$ are all of

the form $1+A^0$. If A^{-1} is unique, there is no A^0 (since $A^{-1}+A^0$ would be another rh. reciprocal), and thus $X = 1$ is the only solution of $AX = A$. Now, if $AA^{-1}.A$ is associative, $(AA^{-1})A = A(A^{-1}.A) = A$, so that $A^{-1}.A = 1$. Moreover, if B is a lh. reciprocal such that BAA^{-1} is associative, $B = B(AA^{-1}) = (BA)A^{-1} = A^{-1}$.

(ii) From ${}^{-1}AA = 1$ and $AA^{-1} = 1$, we get

$$({}^{-1}AA)A^{-1} = A^{-1}, \qquad {}^{-1}A(AA^{-1}) = {}^{-1}A,$$

so that, when ${}^{-1}AAA^{-1}$ is associative, ${}^{-1}A = A^{-1}$. Fixing A^{-1}, say, every ${}^{-1}A$ is equal to this fixed A^{-1}, i.e. there is one ${}^{-1}A$ only, and similarly for A^{-1}. It may happen, though, that A has no A^{-1} and has infinitely many ${}^{-1}A$.

It follows from I that if, in an associative field F, A has a unique A^{-1}, this A^{-1} is a two-sided reciprocal of A, and unique in F. Or, if, in an F, A has both an A^{-1} and an ${}^{-1}A$, they coincide and form the unique (two-sided) reciprocal of A in F.

II. *If A has an A^{-1} (${}^{-1}A$), and if $\bar{A}$ is a matrix obtained by exchanging certain rows (columns) in A, the corresponding exchange of columns in A^{-1} (rows in ${}^{-1}A$) leads to an $\bar{A}^{-1}$ (${}^{-1}\bar{A}$). If A has an A^0 (${}^0A$), the same matrix is an $\bar{A}^0$ (${}^0\bar{A}$).*

Proof. By hypothesis, $\sum_i a_{ni}x_{ik} = \delta_k^n$ has a solution. Exchange the rows of A by renumbering them, i.e. by establishing a one-one correspondence between n and ν, say, indicated by $\nu = f(n)$, $n = g(\nu)$ from n to ν and from ν to n respectively. The renumbered equations are

$$\sum_i a_{f(n)i}x_{ik} = \delta_k^{f(n)}, \qquad (a_{f(n)i}) = \bar{A}. \tag{3}$$

Now, renumber the unknowns in

$$\bar{A}X' = 1, \text{ i.e. } \sum_i a_{f(n)i}x'_{ik} = \delta_k^n,$$

by rewriting these equations in the form

$$\sum_i a_{f(n)i}x'_{ig(k)} = \delta_{g(k)}^n. \tag{4}$$

When $f(n) = k$, then $n = g(k)$ and thus, for $x'_{ig(k)} = x_{ik}$, (4) coincides with (3), i.e. $x'_{ig(k)} = x_{ik}$ is a solution of $\bar{A}X' = 1$.

This result extends to $AX = B$ in the form: if (1) has a solution x_{ik}, then $x_{ig(k)}$ is a solution of $(a_{f(n)k})X = (b_{f(n)g(k)})$.

The second statement in II follows from the fact that, in $AX = 0$, every column x_{ik}, $i = 1, 2, \ldots, \infty$, satisfies the same equations.

We also notice that if $\bar{A}$ is a matrix obtained by ruling out certain

rows of A, every solution of (1) is a solution of $\bar{A}X = \bar{B}$, where $\bar{B}$ is B without the rows corresponding to the missing rows in A. In particular, if A has an A^{-1}, $\bar{A}$ has an $\bar{A}^{-1}$, and if A has an ${}^{-1}A$, then, leaving out columns in A, we obtain matrices having a lh. reciprocal. It follows, for instance, that if A has no A^{-1}, no addition of new rows leads to a matrix with a rh. reciprocal.

III. *If A and B have rh. (lh.) reciprocals such that $ABB^{-1}.A^{-1}$ $({}^{-1}B.{}^{-1}AAB)$ is associative, $B^{-1}.A^{-1}$ $({}^{-1}B.{}^{-1}A)$ is a rh. (lh.) reciprocal of AB. If there is an ${}^{-1}A$ and a $(AB)^{-1}$ such that ${}^{-1}AAB(AB)^{-1}.A$ is associative, $(AB)^{-1}.A$ is a B^{-1}.*

PROOF. $(AB)B^{-1}.A^{-1} = A(BB^{-1})A^{-1} = 1$. If now we multiply $AB(AB)^{-1} = 1$ by ${}^{-1}A$ on the left and by A on the right, we obtain $B[(AB)^{-1}.A)] = 1$.

We notice that when A and B belong to the same associative field F, and A^{-1} and B^{-1} are unique on the right, $B^{-1}.A^{-1}$ is the unique $(AB)^{-1}$ in F. In fact, in this case, A^{-1} and B^{-1} are also lh. reciprocals and thus it follows from $(AB)X = 1$ that $X = {}^{-1}B.{}^{-1}A$.

IV. *If A^{-1} and $(1+BA^{-1})^{-1}$ exist and $(A+B)A^{-1}(1+BA^{-1})^{-1}$ is associative, $A^{-1}.(1+BA^{-1})^{-1}$ is an $(A+B)^{-1}$. If ${}^{-1}A$, ${}^{-1}(1+{}^{-1}AB)$ exist, and ${}^{-1}(1+{}^{-1}AB).{}^{-1}A(A+B)$ is associative, ${}^{-1}(1+{}^{-1}AB).{}^{-1}A$ is an ${}^{-1}(A+B)$.*

In these statements A and B may be exchanged.

PROOF. $(A+B)A^{-1}.(1+BA^{-1})^{-1} = (1+BA^{-1})(1+BA^{-1})^{-1} = 1$.

We notice that if A^{-1}, B^{-1}, $(1+AB^{-1})^{-1}$, and $(1+BA^{-1})^{-1}$ are two-sided in an F, then

$$B^{-1}.(1+AB^{-1})^{-1} = A^{-1}.(1+BA^{-1})^{-1}.$$

In fact, from $(1+AB^{-1})B = (1+BA^{-1})A = A+B$, we have

$$1+AB^{-1} = (1+BA^{-1})AB^{-1}, \qquad (1+BA^{-1})^{-1}.(1+AB^{-1}) = AB^{-1},$$

$$A^{-1}.(1+BA^{-1})^{-1}.(1+AB^{-1}) = B^{-1}, \text{ etc.}$$

We notice that exchanging the first two rows is the same as multiplying A on the left by the matrix

$$\begin{matrix} 0 & 1 & 0 & 0 & . & . & . \\ 1 & 0 & 0 & 0 & . & . & . \\ 0 & 0 & 1 & 0 & . & . & . \\ 0 & 0 & 0 & 1 & . & . & . \\ . & . & . & . & 1 & . & . \end{matrix}$$

and we rule out the second row, say, of A if we multiply it by the matrix

$$\begin{matrix} 1 & 0 & 0 & 0 & . & . & . \\ 0 & 0 & 1 & 0 & . & . & . \\ 0 & 0 & 0 & 1 & . & . & . \\ . & . & . & . & 1 & . & . \end{matrix}$$

In this way II can be established, at least in simple cases, by III. I owe this remark to Professor A. C. Dixon.

4. Examples. Applications. We shall now give some matrices having no reciprocal, and some having some kind of reciprocal.

If A has a 0 row (column), A has no rh. (lh.) reciprocal, since $\sum_i a_{ni} x_{ik} = \delta_k^n$ reduces to $0 = 1$ when $k = n'$ for which $a_{n'i} = 0$, $i = 1, 2, ..., \infty$.

If the terms of a row (column) in A are the same multiple of another row (column), A has no rh. (lh.) reciprocal, since if $a_{2k} = ca_{1k}$, say, $\sum_i a_{2i} x_{ik} = \delta_k^2$ becomes $c\delta_k^1 = \delta_k^2$, i.e. $0 = 1$ when $k = 2$.

More generally, if $\sum_{j=1}^{\infty} c_j a_{ji} = 0$ and $c_1 \neq 0$, say, then $\sum_i a_{1i} x_{ik} = \delta_k^1$ becomes $-\frac{1}{c_1} \sum_{j=2}^{\infty} c_j \delta_k^j = \delta_k^1$, i.e. $0 = 1$, when $k = 1$, provided the two summations in $\sum_{i=1}^{\infty} \sum_{j=2}^{\infty} c_j a_{ji} x_{ik}$ can be interchanged. This is the case if the second sum is finite, or if c_j considered as a diagonal matrix, a_{jk} and x_{ki} all belong to the same associative field. In an associative field with a regular bound, this condition only means that $|c_j|$ is bounded.

We say that A is a *cross-matrix*, if all its non-vanishing terms are in a finite number of rows *and* columns, as, for instance, row-bounded or column-bounded matrices.

No cross-matrix has a reciprocal.

PROOF. By hypothesis, there is an r such that $a_{r+n,r+k} = 0$ for every n and k. Take $r+1$ rows such that $a_{ni} = 0$ for $i > r$, and take them for the first $r+1$ rows. Let Δ denote the determinant (a_{nk}) $(n, k = 1, 2, ..., r)$. If $\Delta = 0$, there is a linear relation between the r linear forms $\sum_{i=1}^{r} a_{ni} x_{ik}$ $(n = 1, 2, ..., r)$, leading to $\sum_{i=1}^{r} c_i \delta_k^i = 0$, which is contradictory for $k = j$ if $c_j \neq 0$ (and there is such a c_j unless all the a_{ni} vanish). If $\Delta \neq 0$, the equations $\sum_{i=1}^{r} a_{ni} x_{ik} = \delta_k^n$ $(n = 1, 2, ..., r)$ are equivalent to $\Delta . x_{nk} = \Delta_{nk}$, where Δ_{nk} is Δ in which the nth

column has been replaced by δ_k^n (Cramer's rule). Since $\Delta_{n,r+1} = 0$, it follows that $x_{n,r+i} = 0$ for $i > 0$, and thus $\sum_{i=1}^{r} a_{r+1,i} x_{ik} = \delta_k^{r+1}$ reduces to $0 = 1$ for $k = r+1$.

In case of row-finite matrices, see Toeplitz (1909), $AX = 1$ reduces to

$$\sum_{i=1}^{r_n} a_{ni} x_{ik} = \delta_k^n,$$

so that in many cases we can solve these equations one by one or in groups. For example, if the rows are of unequal length and there is no zero row, we can arrange them in increasing order to form a stair-like matrix. We can also dispose of the arbitrary terms to obtain a row-finite reciprocal.

In particular, for a lower (upper) semi-matrix where $a_{nn} \neq 0$, we obtain a unique rh. (lh.) reciprocal which is a lower (upper) semi-matrix with $a_{nn}^{-1} = 1/a_{nn}$. It follows that this rh. (lh.) reciprocal is at the same time a lh. (rh.) reciprocal and is the only two-sided reciprocal among lower (upper) semi-matrices.

The same lower (upper) semi-matrix A, however, may still have infinitely many lh. (rh.) reciprocals. For example, if the non-vanishing terms are $a_{k,k-1}$, a_{kk} (two diagonals), $XA = 1$ reduces to

$$x_{n,k+1} a_{k+1,k} + x_{nk} a_{kk} = \delta_k^n,$$

which, by fixing n and varying k, determines the nth row of X when the first column of X is arbitrarily given.

If the matrix has no special form leading to simple equations for its reciprocal, we have to restrict ourselves to an associative field with a bound. We say that $\sum_{p=1}^{\infty} A^{(p)}$ is (absolutely) convergent, if $\sum_{p=1}^{\infty} a_{nk}^{(p)}$ is (absolutely) convergent for every n and k.

I. *If A is a matrix of a bounded associative field F and if $|A| < 1$, then $1 + \sum_{p=1}^{\infty} (-1)^p A^p$ is absolutely convergent, its sum is the unique, and two-sided, reciprocal of $1+A$ in F, and*

$$\left|1 + \sum_{p=1}^{\infty} (-1)^p A^p\right| \leqslant \frac{1}{1-|A|}. \tag{1}$$

For regular bounds such that $|A| < \frac{1}{2}$, we also have

$$\left|1 + \sum_{p=1}^{\infty} (-1)^p A^p\right| > \frac{1}{1+|A|}. \tag{2}$$

PROOF. It follows from $|A^p| \leqslant |A|^p$ that, if $\sum_{p=1}^{\infty} |c_p||A|^p = c$ and $\sum_{p=1}^{\infty} |g_p||B|^p = g$ are convergent, $C = \sum_p c_p A^p$ and $G = \sum_p g_p B^p$ are matrices of F such that $|C| \leqslant c$, $|G| \leqslant g$. Moreover, it follows from

$$|(CG)_{nk}| \leqslant \sum_i \sum_p |c_p||(A^p)_{ni}| \sum_q |g_q||(B^q)_{ik}| \leqslant cg, \tag{3}$$

that the terms of the product $\sum c_p A^p \sum g_q B^q$ can be rearranged at will, without changing its sum. Applying this remark to $C = 1+A$ and $G = 1+\sum_{p=1}^{\infty} (-1)^p A^p$, we readily prove the first statement.

Also, for regular bounds, it follows from $|A| < 1$ that

$$|1 \pm A| \geqslant |(1 \pm A)_{nk}| \geqslant 1 - |a_{nk}| \geqslant 1 - |A|.$$

Thus, putting $B = (1+A)^{-1}$, we have

$$|AB| \leqslant |A||B| \leqslant \frac{|A|}{1-|A|} < 1, \quad \text{if } |A| < \tfrac{1}{2}.$$

Since B satisfies the equation $B = 1 - AB$, we have

$$|B| \geqslant 1 - |AB| > 1 - |A||B| > 0,$$

and thus

$$|B| > \frac{1}{1+|A|}.$$

An immediate extension of I, by 3. IV, is

II. *If, in a bounded associative field F, an A^{-1} exists such that $|BA^{-1}| < 1$, then $A^{-1}\left[1+\sum_{p=1}^{\infty}(-1)^p(BA^{-1})^p\right]$ is a rh. reciprocal of $A+B$ in F. If A^{-1} is the only rh. reciprocal of A in F and $|BA^{-1}| < 1$, then $A^{-1}\left[1+\sum_{p=1}^{\infty}(-1)^p(BA^{-1})^p\right]$ is the unique, and two-sided, rh. reciprocal of $A+B$ in F.*

For H matrices, see Riesz, p. 113.

If the lower bound of $|d_n|$ is $d > 0$, D has a unique reciprocal and any regular bound of D^{-1} is $1/d$. Hence, if $|B| < d$ (regular bound), $|BD^{-1}| \leqslant |B||D^{-1}| < 1$ and II proves that $D+B$ has a unique and two-sided reciprocal in F containing both D and B. For matrices which are K_r and K_c matrices at the same time, this result has been recently proved by R. G. Cooke (1931). It follows for an associative field F with a regular bound that if $0 < d \leqslant a_{nn} \leqslant d'$ and $|A^{(0)}| < d$, where $A^{(0)}$ is A in which the diagonal terms have been replaced by 0, then A has a unique reciprocal in F.

Putting $x_{k1} = x_k$, $x_{ki} = 0$ if $i > 1$, $b_{n1} = b_n$, $b_{ni} = 0$ if $i > 1$, we can write

$$\sum_{k=1}^{\infty} a_{nk} x_k = b_n \tag{4}$$

in the matrix form

$$AX = B \tag{5}$$

with the proviso that we consider only solutions of (5) in which all the terms vanish beyond the first column. It follows that, if A has an A^{-1} such that $AA^{-1}.B$ is associative, $A^{-1}.B$ is a solution of (4), since with B also $A^{-1}.B$ is a first-column matrix. Moreover, if A^{-1} is unique in an associative field F, $A^{-1}.B$ is the unique solution of (4) in F.

For example, if A is a K_r matrix with a K_r reciprocal A^{-1}, and if the sequence $|b_n|$ is bounded, $A^{-1}.B$ exists and $x_k = (A^{-1}.B)_{k1}$ is a solution of (4) such that $|x_k|$ is bounded. If A is a K_c matrix with a K_c reciprocal A^{-1}, and if $\sum_n |b_n|$ is convergent, $x_k = (A^{-1}.B)_{k1}$ is a solution of (4) such that $\sum_k |x_k|$ is convergent.

These remarks apply in particular if A, D, and B belong to an associative field F with a regular bound, and if $|A - D|$ is less than the lower bound of $|d_n|$. If A is a K_r matrix, $D = 1$, $|b_n|$ bounded, we obtain a slight extension of Dixon's first result (Dixon (1902), p. 191).

We close this article by some remarks tending to show that, in infinite systems of equations symbolized by

$$AX = B, \tag{6}$$

the transformation of A into a diagonal matrix $D = \Omega^{-1}.A\Omega$ corresponds to the elimination process.

If A, B, C are in F, then (6) and $AXC = BC$ or $CAX = CB$ are equivalent in F, i.e. have the same solutions in F provided that C has in F a rh. or a lh. reciprocal respectively. Therefore, if there is an Ω^{-1} in the field F containing A, B, and Ω, in which case Ω^{-1} has a lh. reciprocal in F, viz. Ω, then (6) and

$$\Omega^{-1}.A\Omega.\Omega^{-1}.X\Omega = \Omega^{-1}.B\Omega, \quad \text{i.e. } DY = \bar{B} \tag{7}$$

are equivalent.

Now, if no d_n vanishes, the only solution of (7) in F is $Y = D^{-1}.\bar{B}$ and thus the only solution of (6) in F is

$$X = \Omega D^{-1}.\bar{B}\Omega^{-1} = \Omega D^{-1}.\Omega^{-1}.B.$$

If d_n vanishes for the suffixes n', say, the necessary and sufficient condition that (7), and thus also (6), have a solution in F is that $b_{n'k} = 0$ for every k, and in this case the rows $x_{n'k}$ are arbitrary.

Since $\Omega^{-1}.A\Omega = D$ is a diagonal matrix, $(\Omega^{-1}.A\Omega)_{nk} = 0$ unless $k = n$. Therefore, if $d_{n'} = 0$, we have

$$(\Omega^{-1}.A\Omega)_{n'k} = 0 \text{ for every } k. \tag{8}$$

On the other hand, the explicit form of the conditions $\bar{b}_{n'k} = 0$ is

$$(\Omega^{-1}.B\Omega)_{n'k} = 0 \text{ for every } k, \tag{9}$$

i.e. *when $d_n = 0$ for some affixes, A and B have to satisfy the same conditions in order that* (6) *should have a solution.* This is the extension of a well-known property of finite systems to a class of infinite systems of simultaneous equations.

An obvious generalization of this result is as follows. If (i) A, B, Ω, W belong to the same associative field F, (ii) Ω has a rh., W a lh. reciprocal in F, (iii) $WA\Omega = D$, then (6) and

$$WA\Omega.\Omega^{-1}.X\Omega = WB\Omega, \text{ i.e. } DY = \bar{B}, \tag{10}$$

are equivalent and the same conclusion holds as before with conditions (8) and (9) replaced by

$$(WA\Omega)_{n'k} = 0, \qquad (WB\Omega)_{n'k} = 0, \quad \text{for every } k. \tag{11}$$

5. Two general types of linear equations. If X is a solution of the linear equation

$$AXA' + BXB' = C, \tag{1}$$

every solution is of the form $X+\Omega$, where Ω is a solution of

$$AXA' + BXB' = 0. \tag{2}$$

Multiplying by B^{-1} on the left and by A'^{-1} on the right, (1) and (2) assume the simpler forms

$$AX + XB = C, \tag{3}$$

$$AX + XB = 0. \tag{4}$$

For (4) we have the following results. If Ω is a solution of (4) and C any matrix commutable with A ($AC = CA$) such that $AC\Omega$ and $C\Omega B$ are associative, then $C\Omega$ is another solution of (4). In fact

$$AC\Omega + C\Omega B = C(A\Omega + \Omega B) = 0.$$

In a like manner ΩC is a solution if $CB = BC$.

Conversely, if $C\Omega$ is another solution of (4), i.e. if $A\Omega + \Omega B = 0$ and $AC\Omega + C\Omega B = 0$ and the products are associative, we have $AC\Omega = CA\Omega$. Thus, when there is an Ω^{-1}, $CA = AC$. We finally remark that if there is an $^{-1}\Omega$, and if $W.^{-1}\Omega\Omega$ is associative, W can be written in the form $W = C\Omega$, for $W = X\Omega$ can be solved. Therefore, in that case, all the solutions of (4) are of the form $C\Omega$, where $CA = AC$. This is the case in an associative field F. Hence

I. *If A and B are in an associative field F and Ω is a solution of* (4) *in F possessing a two-sided Ω^{-1} in F, all the solutions of* (4) *in F are of the form $C\Omega$, where C is any matrix in F commutable with A, and all such products are solutions of* (4).

Suppose now that A, B, Ω are in F. If there is a B^{-1} and an Ω^{-1} or an ${}^{-1}\Omega$ in F, then $-\Omega B^{-1}.\Omega^{-1}$, or $-\Omega B^{-1}.{}^{-1}\Omega$, is an A^{-1} in F. Therefore, if A has no A^{-1} in F and B has a B^{-1} in F, (4) cannot be solved by an Ω in F having a reciprocal in F. In this statement A and B can obviously be exchanged.

In some cases a solution of (3) can be obtained by successive approximations as shown in

II. *Put $A = D+A^{(0)}$, $B = \Theta+B^{(0)}$, $C = \Gamma+C^{(0)}$, where $d_n = a_{nn}$, $\theta_n = b_{nn}$, $\gamma_n = c_{nn}$, and suppose that* (i) $|d_n-\theta_k| \geqslant d > 0$, (ii) $|C|$ *exists and $|A^{(0)}|+|B^{(0)}| = e < d$, where the bars denote a regular bound. Then*

$$X = \Delta+\sum_{s=1}^{\infty} X^{(s)} \tag{5}$$

is a solution of

$$AX-XB = C, \tag{6}$$

where Δ, $X^{(1)}$, $X^{(2)}$,... are successively determined by the equations

$$\left.\begin{array}{llll} & (D-\Theta)\Delta = \Gamma & & \\ DX^{(1)} & -X^{(1)}\Theta & = C^{(0)}+\Delta B^{(0)} & -A^{(0)}\Delta \\ DX^{(2)} & -X^{(2)}\Theta & = X^{(1)}B^{(0)} & -A^{(0)}X^{(1)} \\ \cdot\ \cdot\ \cdot & \cdot\ \cdot\ \cdot & \cdot\ \cdot\ \cdot & \cdot\ \cdot\ \cdot \\ DX^{(s+1)} & -X^{(s+1)}\Theta & = X^{(s)}B^{(0)} & -A^{(0)}X^{(s)} \\ \cdot\ \cdot\ \cdot & \cdot\ \cdot\ \cdot & \cdot\ \cdot\ \cdot & \cdot\ \cdot\ \cdot \end{array}\right\}. \tag{7}$$

Moreover,

$$|X| \leqslant \frac{|\Gamma|+|C^{(0)}|}{d-e}. \tag{8}$$

Proof. $(7)_1$ gives $\delta_n = \gamma_n/(d_n-\theta_n)$. From $(7)_2$

$$x_{nk} = \frac{c_{nk}^{(0)}+\delta_n b_{nk}^{(0)}-a_{nk}^{(0)}d_k}{d_n-\theta_k}, \qquad |X^{(1)}| \leqslant \frac{|C^{(0)}|+|\Delta|e}{d}.$$

From the general equation (7),

$$x_{nk}^{(s+1)} = \frac{\sum x_{ni}^{(s)} b_{ik}^{(0)} - \sum a_{ni}^{(0)} x_{ik}^{(s)}}{d_n-\theta_k},$$

$$|X^{(s+1)}| \leqslant |X^{(s)}|\frac{e}{d} \leqslant \frac{|C^{(0)}|}{d}\left(\frac{e}{d}\right)^s+|\Delta|\left(\frac{e}{d}\right)^{s+1}. \tag{9}$$

Therefore

$$|X| \leqslant |C^{(0)}|\frac{1}{d-e}+|\Delta|\frac{d}{d-e}, \qquad \left|\sum_{s=1}^{\infty} X^{(s)}\right| \leqslant \frac{|C^{(0)}|}{d-e}+|\Delta|\frac{e}{d-e}, \tag{10}$$

and from the first equation

$$|\Delta| \leqslant |\Gamma| d^{-1}.$$

We may add that X will have a reciprocal in the field when, for instance,

$$d(|C^{(0)}|+\delta' e) < \gamma(d-e), \tag{11}$$

where γ is the lower bound of $|\gamma_n|$, δ' is the upper bound of $|\delta_n|$.

If $\Gamma = 0$ this method gives $\Delta = 0$; if $C = 0$ (the homogeneous case), the method leads to the trivial solution $X = 0$, for $\Delta = 0$. However, if $\Theta = D$, the first equation is satisfied for every Δ. In this case, for the transformation of matrices into diagonal matrices, we may make use of the following remark, which I owe to Dr. R. G. Cooke.

If X is the solution of

$$AX-XD = C,$$

and if C, ^{-1}X, X are associative, then

$$(A-C.{}^{-1}X)X-XD = 0,$$

and thus X transforms $A-C.{}^{-1}X$ into D. Some examples also constructed by Dr. Cooke tend to show the efficiency of this procedure.

Our method fails, however, for

$$AX-XA = 1 \tag{12}$$

which, in particular cases, can be solved step by step. For example, if the non-vanishing terms of A are $a_{n,n+1}$, we have

$$a_{n,n+1}x_{n+1,1} = \delta_1^n, \qquad a_{n,n+1}x_{n+1,k}-x_{n,k-1}a_{k-1,k} = \delta_k^n \quad (k > 1).$$

These equations leave the first row of X arbitrary; the other terms of the kth column are

$$x_{2k} = \frac{x_{1,k-1}a_{k-1,k}}{a_{12}}, \qquad x_{3k} = \frac{x_{2,k-1}a_{k-1,k}}{a_{23}}, \ldots, x_{k-1,k} = \frac{x_{k-2,k-1}a_{k-1,k}}{a_{k-2,k-1}},$$

$$x_{kk} = x_{k-1,k-1} = \ldots = x_{11}, \qquad x_{k+1,k} = \frac{k}{a_{k,k+1}}, \; x_{k+i,k} = 0 \quad (i > 1).$$

These relations determine the kth column in terms of the elements of the $(k-1)$th column. Finally, $x_{21} = 1/a_{12}$, $x_{n1} = 0$ $(n > 2)$.

To show the peculiar nature of (12) we first make the obvious remark that (12) has no diagonal solution since $a_{nk}\delta_k-\delta_n a_{nk} = \delta_k^n$ is contradictory for $n = k$. It follows that (12) has no solution which is the transform of a diagonal matrix, since on multiplying

$$A\Omega^{-1}D\Omega-\Omega^{-1}D\Omega A = 1$$

by Ω on the left and by Ω^{-1} on the right, we obtain

$$\Omega A\Omega^{-1}D-D\Omega A\Omega^{-1} = 1,$$

so that D would be the solution of (12) where A is replaced by its transform.

Another less obvious property of (12) is the following. If B satisfies (12), B^n satisfies

$$AB^n - B^nA = nB^{n-1} \tag{13}$$

provided A and B belong to the same associative field F, with a regular bound, say. Now, take matrices A_ν from F such that

$$\overline{\lim}\, \sqrt[\nu]{|A_\nu|} = a^{-1}, \qquad |B| < a,$$

and consider the matrix function

$$F(X) = \sum_{\nu=0}^{\infty} A_\nu X^\nu. \tag{14}$$

It follows from (13) that, if A_ν are commutable with A,

$$AF(B) - F(B)A = F'(B), \tag{15}$$

where $F'(X)$ is the formal derivative of $F(X)$, i.e.

$$F'(X) = \sum_{\nu=1}^{\infty} \nu A_\nu X^{\nu-1}. \tag{16}$$

Similarly, if $f(X) = \sum_{\nu=1}^{\infty} C_\nu (X-B)^\nu$, we have

$$f'(X) = \sum_{\nu=1}^{\infty} \nu C_\nu (X-B)^{\nu-1},$$

from which it follows that

$$C_\nu = f^{(\nu)}(B)/\nu!. \tag{17}$$

Therefore, if we expand $F(X) = \sum A_\nu X^\nu$ about B, i.e. if we replace X by $(X-B)+B$, use the binomial theorem, and rearrange the series in powers of $X-B$, we obtain

$$F(X) = \sum C_\nu (X-B)^\nu. \tag{18}$$

This process implies that we restrict ourselves to matrices X which are commutable with B, and for which the resulting series (18) is absolutely convergent at least if $|X-B| < a-|B|$.

If $|B| < 2a$, i.e. if $X = 0$ is in the circle of convergence of (18), we can re-transform (18) into (14) and thus obtain

$$A_\nu = C_\nu - C_\nu^{\nu+1} C_{\nu+1} B + C_\nu^{\nu+2} C_{\nu+2} B^2 - \ldots. \tag{19}$$

This shows that, if $C_\nu = 0$ for every ν, $A_\nu = 0$ for every ν. Now, if B satisfies (12), every absolutely convergent power series with matrix coefficients commutable with A satisfies (15). Thus also $F'(B)$, $F''(B)$, etc., satisfy it. Therefore, if $F(B) = 0$, we have $F'(B) = 0$, $F''(B) = 0$, etc., i.e. $C_\nu = 0$ for every ν, and thus also $A_\nu = 0$, i.e. *B is not a root of an equation like $F(B) = 0$.*

Analytic continuation leads to some obvious extensions of this result. A consequence of our result is that (12), where A is a finite

matrix, has no finite matrix solution. In fact, every finite matrix satisfies the so-called rank equation.

BIBLIOGRAPHY

M. Bôcher and L. Brand, 'On linear equations with an infinite number of variables': *Ann. of Math.* (2) 13 (1912), 167–86.

R. G. Cooke, 'A theorem on reciprocals': *Journal London Math. Soc.* 6 (1931), 269–72.

P. Dienes, 'The exponential function in linear algebras': *Quart. J. of Math.* (Oxford), 1 (1930), 300–9.

P. Dienes, *The Taylor Series*, Oxford University Press (1931), chap. xii, 384–436.

A. C. Dixon, 'On a class of matrices of infinite order and on the existence of "matricial" functions on a Riemann surface': *Trans. of Cambridge Phil. Soc.* 19 (1902), 190–233.

H. Hahn, 'Über die Integrale des Herrn Hellinger und die Orthogonalinvarianten der quadratischen Formen von unendlichvielen Veränderlichen': *Monatshefte f. Math. und Phys.* 23 (1912), 161–224.

H. Hahn, 'Über Folgen linearer Operationen': ibid. 32 (1922), 3–88.

E. Hellinger, 'Neue Begründung der Theorie quadratischer Formen von unendlichvielen Veränderlichen': *Crelle*, 136 (1909), 210–71.

E. Hellinger und O. Toeplitz, 'Grundlagen für eine Theorie der unendlichen Matrizen': *Göttinger Nachrichten* (1906), pp. 351–5.

E. Hellinger und O. Toeplitz, 'Grundlagen für eine Theorie der unendlichen Matrizen': *Math. Annalen*, 69 (1910), 289–330.

E. Helly, 'Über Systeme linearer Gleichungen mit unendlich vielen Unbekannten': *Monatshefte f. Math. und Phys.* 31 (1921), 60–91.

E. Hilb, 'Über die Auflösung von Gleichungen mit unendlich vielen Unbekannten': *Sitz.-ber. d. Phys.-Med. Soc., Erlangen* (1908), 84–9.

E. Hilb, 'Über die Auflösung unendlich vieler linearer Gleichungen mit unendlich vielen Unbekannten': *Math. Annalen*, 70 (1910), 79–86.

D. Hilbert, *Grundzüge einer allgemeinen Theorie der linearen Integralgleichungen*, Teubner, Leipzig, 1912.

D. Hilbert, 'Wesen und Ziele einer Analysis der unendlichvielen unabhängigen Variabeln': *Rend. di Palermo*, 27 (1909), 59–74.

J. Hyslop, 'A contribution to the theory of bounded matrices and quadratic forms with an infinity of variables': *Proc. London Math. Soc.* (2) 24 (1926), 264–304.

G. Kowalewski, *Einführung in die Determinantentheorie*, Veit, Leipzig, 1909.

F. Riesz, *Les Systèmes d'équations linéaires à une infinité d'inconnues*. Paris, 1913.

E. Schmidt, 'Über die Auflösung linearer Gleichungen mit abzählbar unendlich vielen Unbekannten': *Rend. di Palermo*, 24 (1908), 53–77.

O. Toeplitz, 'Über die Auflösung unendlichvieler linearer Gleichungen mit unendlichvielen Unbekannten': *Rend. di Palermo*, 28 (1909), 88–96.

O. Toeplitz, 'Zur Theorie d. quadratischen u. bilinearen Formen von unendlichvielen Veränderlichen': *Math. Annalen*, 70 (1911), 351–76.

J. L. Walsh, 'On the solution of linear equations in infinitely many variables by successive approximations': *American J. of Math.* 42 (1920), 91–6.

ON CERTAIN INTEGRAL EQUATIONS

By E. HOPF (*Boston, U.S.A.*)

[Received 4 August 1932]

1. THE present note contains the proof of the following:

THEOREM. *Let $K(t)$ $(0 < t < \infty)$ be a real, piece-wise continuous and not identically vanishing function with the properties*

$$(a) \qquad \int_0^\infty \{K(t)e^{-\theta t}\}^2\,dt < \infty \qquad \text{(every } \theta < 1\text{)},$$

$$(b) \qquad K(t) \geqslant 0.$$

We set $K_{n+1}(t) = \int_t^\infty K_n\,dt$, $K_1 = K$. *Furthermore we suppose that*

$$(c) \qquad K_2(0) = \tfrac{1}{2}.$$

The integral equation

$$f(\tau) = \int_0^\infty K(|\tau - t|)f(t)\,dt \tag{1}$$

possesses then one and only one solution satisfying

$$\frac{f(\tau)}{\tau} \to 1, \qquad \tau \to \infty. \tag{2}$$

Its value at $\tau = 0$ *is*

$$f(0) = +\sqrt{\{2K_4(0)\}}. \tag{3}$$

For the validity of (3), (*c*) is the most essential part of the hypothesis. No generalization could be found in the case $2K_2(0) \neq 1$.

The particular case $K(t) = \frac{1}{2}\int_1^\infty e^{-\lambda t}\lambda^{-1}\,d\lambda$ is of interest in modern astrophysics: (*c*) is easily seen to be fulfilled. The relation (3), $f(0) = 1/\sqrt{3}$, yields, then, the ratio of the boundary temperature to the effective temperature for a star in radiative equilibrium,†

$$\left(\frac{T_0}{T_e}\right)^4 = \frac{\sqrt{3}}{4}.$$

Under the hypothesis (*a*), a theory of (1) has been given by N. Wiener and the author.‡ The discussion of (1) is governed by the

† M. Bronstein, *Zeits. für Physik*, 59 (1929), 144; E. Hopf, *M.N. R.A.S.* 90 (1930), 287.

‡ N. Wiener und E. Hopf, 'Über eine Klasse singulärer Integralgleichungen': *Berlin Sitzungsber. der Akad.*, 1931.

characteristic equation

$$\kappa(z) \equiv \int_0^\infty K(t)(e^{zt}-e^{-zt})\,dt = 1. \tag{4}$$

Here $\kappa(z) = \kappa(x+iy)$ is regular in the strip $|x| < 1$ and tends to zero as $|y| \to \infty$, the convergence being uniform in every partial strip $|x| \leqslant \theta < 1$.

2. General facts concerning (1)

In order to determine all solutions of (1) with the property

$$f(\tau) = o(e^{\theta\tau}) \qquad (0 < \theta < 1), \tag{5}$$

we proceed as follows. Let $z_1, z_2, \ldots, z_{2n}$ be the roots of (4) lying in the strip $|x| < \theta$. We set

$$P(z) = \prod_1^{2n} (z-z_\nu)$$

and

$$h(z) = \log\left[\{1-\kappa(z)\}\frac{(z^2-1)^n}{P(z)}\right]. \tag{6}$$

The function between the brackets does not vanish in $|x| < \theta$ and is positive for $y = 0$. The logarithm is chosen to be real for real z. Then $h(z) \to 0$ for $(y) \to \infty$, $|x| < \theta$; $|h(x+iy)|$ is a quadratically summable function of y.

All solutions of (1) satisfying (5) are contained in the formulae

$$f(\tau) = \frac{1}{2\pi i}\int_{-\theta-i\infty}^{-\theta+i\infty} F(z)e^{-\tau z}\,dz, \tag{7}$$

$$F(z) = \frac{(z-1)^n Q_{n-1}(z)}{P(z)}\exp\left[-\frac{1}{2\pi i}\int_{\theta-i\infty}^{\theta+i\infty}\frac{h(\zeta)}{\zeta-z}\,d\zeta\right], \tag{8}$$

where Q_{n-1} is an arbitrary polynomial of the indicated degree. The integral (7) is to be understood in the sense of the Plancherel theory of Fourier integrals. Furthermore, each $f(\tau)$ given by (7), (8) vanishes for $\tau < 0$ and is continuous for $\tau \geqslant 0$; (7) can be inverted, and so

$$F(z) = \int_0^\infty f(t)e^{zt}\,dt \qquad (x < 0). \tag{9}$$

From (7), (8) the asymptotic form of the solutions can be obtained,† for τ large,

$$f(\tau) = \sum Q^*(\tau)e^{-z^*t} + o(e^{\theta\tau}),$$

† By transferring the integration abscissa in (7) to the right as far as θ and by taking account of the residues of $F(z)e^{-z\tau}$.

where z^* is a root of $P(z)$, and where Q^* is a certain polynomial of degree less than the multiplicity of z^*.

3. Proof of the theorem

We have $\kappa(0) = 2K_2(0)$: (c) means therefore that $z = 0$ is a root of (4). Since $\kappa(z)$ is an even function, we have $\kappa'(0) = 0$. Furthermore,

$$\kappa''(0) = 2\int_0^\infty t^2K(t)\,dt = 4K_4(0) > 0. \tag{10}$$

Therefore $z = 0$ is a double root of (4), and, according to (b), no other root can lie on the imaginary axis $x = 0$. The roots of (4) can thus be arranged in the form

$$z_1,\ldots,z_{n-1},0,0,-z_1,\ldots,-z_{n-1},$$

where
$$-\theta < x_\nu = \mathrm{R}(z_\nu) < 0. \tag{11}$$

On writing
$$P^*(z) = \prod_1^{n-1}(z-z_\nu), \tag{12}$$

we have
$$P(z) = (-1)^{n-1}z^2P^*(z)P^*(-z). \tag{13}$$

If, now, (1) has a solution which is approximately linear for τ large, we infer from (9) that $F(z)$ is regular for $x < 0$. This regularity can be obtained only by setting

$$Q_{n-1}(z) = cP^*(z)$$

in (8). We have thus

$$F(z) = (-1)^n c\frac{(z-1)^n}{z^2P^*(-z)}\exp\left[-\frac{1}{2\pi i}\int_{\theta-i\infty}^{\theta+i\infty}\frac{h(\zeta)}{\zeta-z}\,d\zeta\right]. \tag{14}$$

Conversely, we prove that (7), (14) actually yield an asymptotically linear solution of (1). $F(z)e^{-\tau z}$ can have poles only at $z = 0, -z_1, \ldots, -z_{n-1}$. Only τ and bounded terms can therefore occur in the asymptotic expression for $f(\tau)$. For $c \neq 0$, however, τ cannot be missing, since a bounded $f(\tau)$ would, according to (9), imply

$$|zF(z)| < \text{const. } |z/n| \qquad (x < 0),$$

a contradiction to (14).

For $y = 0$, $x < 0$, we infer from (9) that

$$|x|F(x) = \int_0^\infty f(t/|x|)e^{-t}\,dt,$$

thus yielding
$$|x|F(x) \to f(0) \qquad (x \to -\infty).$$

On the other hand, from (14),

$$|x|F(x) \to c \qquad (x \to -\infty),$$

thus yielding

$$c = f(0). \tag{15}$$

Now we investigate $F(z)$ at $z = 0$. If (2) is satisfied, we know that $f(\tau) = \tau + O(1)$. According to (9) this implies, for $x < 0$, $y = 0$,

$$F(x) = \frac{1}{x^2} + \int_0^\infty O(1)e^{xt}\,dt = \frac{1}{x^2} + O\left(\frac{1}{|x|}\right),$$

thus yielding

$$\lim_{x \to -0} x^2 F(x) = 1. \tag{16}$$

On the other hand, from (14),

$$\lim_{z \to 0} z^2 F(z) = \frac{c}{P^*(0)} \exp\left[-\frac{1}{2\pi i} \int_{\theta - i\infty}^{\theta + i\infty} \frac{h(\xi)}{\xi}\,d\xi\right]. \tag{17}$$

The integral in (17) may be evaluated as follows. The path of integration, $\mathrm{R}(\xi) = \theta$, may be transferred to the left and deformed into another broken line, which consists of the two parts $(-i\infty, -i\rho)$, $(i\rho, i\infty)$ of the imaginary axis and of the right half of the circle $|\xi| = \rho$. It is now essential that the integrals along the first parts cancel each other, for $h(\xi)/\xi$ is an odd function. Hence

$$\int_{\theta - i\infty}^{\theta + i\infty} \frac{h(\xi)}{\xi}\,d\xi = i \int_{-\frac{1}{2}\pi}^{+\frac{1}{2}\pi} h(\rho e^{i\phi})\,d\phi.$$

This tends to $\pi i h(0)$ as $\rho \to 0$. Therefore (17) becomes

$$\lim_{z \to 0} z^2 F(z) = \frac{c}{P^*(0)} e^{-\frac{1}{2}h(0)},$$

thus yielding, according to (15) and (16),

$$f(0) = P^*(0) e^{\frac{1}{2}h(0)}. \tag{18}$$

From (6), (10), and (13), however,

$$h(0) = \log \tfrac{1}{2}\kappa''(0) - 2 \log P^*(0),$$

the logarithm being real, and thus (18) gives the result (3)

ON A PROBLEM IN ADDITIVE ARITHMETIC (II)

By L. CARLITZ (*Cambridge*)

[Received 12 August 1932]

Introduction

In a paper in this journal,† bearing the same title as the present one, I considered the sum

$$\sum_{n=n_1+\ldots+n_\nu} \alpha(n_1)\ldots\alpha(n_\nu), \tag{1}$$

taken over all the partitions of n into ν positive integers, $\alpha(n)$ being a 'factorable' arithmetic function satisfying certain order conditions. By applying the Hardy-Littlewood method, an asymptotic expression was found for (1), at least for $\nu \geqslant 3$. The present paper deals with the somewhat more general sum

$$\sum \alpha_1(n_1)\ldots\alpha_\nu(n_\nu), \tag{2}$$

for all values of $\nu \geqslant 2$, where the functions $\alpha_i(n)$ are similar to $\alpha(n)$ above, that is, they satisfy the same general conditions that are satisfied by $\alpha(n)$. The method used is now quite elementary. It was suggested by Estermann's‡ proof of a theorem due to Evelyn and Linfoot§—a theorem included as a special case of the theorem to be proved about (2) (for $\nu = 2$). We find that the sum in (2) is

$$Bn^{\nu-1}\alpha^{(\nu)}(n)+O(n^{\nu-\frac{1}{2}+\epsilon}), \tag{3}$$

for all positive ϵ; where B is independent of n, but depends on the functions $\alpha_i(n)$; $\alpha^{(\nu)}(n)$ depends on the $\alpha_i(n)$, and is a function of exactly the same kind. While the functional dependence of $\alpha^{(\nu)}$ on α_i is fairly simple, the expression of $\alpha^{(\nu)}$ explicitly in terms of α_i is quite complicated. Accordingly, we postpone the precise statement of the theorem to be proved until we have made the necessary definitions of certain sets of functions derived from the α_i (see § 1).

Instead of considering directly the sum (2), we find it convenient to consider the more general sum

$$\sum_{n=n_1+\ldots+n_\nu} n_1^{h_1}\alpha_1(n_1)\ldots n_\nu^{h_\nu}\alpha_\nu(n_\nu), \tag{4}$$

where the h_i are non-negative integers. The generalized sum (4) is of some interest in itself; the point of the generalization is, however,

† *Quart. J. of Math.* (Oxford), 2 (1931), 97–106. Cited as I.

‡ *Journal London Math. Soc.* 6 (1931), 37–40.

§ *Journal für Math.* 164 (1931), 131–40.

that once the case $\nu = 2$ has been treated, we are able to pass immediately from the theorem for ν to the corresponding theorem for $\nu+1$: furthermore, the presence of the h_i adds very little complication to the treatment of the case $\nu = 2$.

The case $\nu = 2$ of either (2) or (4) naturally suggests the conjugate sum

$$\sum_{n\leq x} \alpha_1(n)\alpha_2(n+k). \tag{5}$$

This sum can be treated by the method employed for the sum (4). If k be fixed, then, for $x \to \infty$, the sum in (5) is

$$Bx\alpha^{(2)}(k)+O(x^{\frac{1}{2}+\epsilon}), \tag{6}$$

where B is free of both k and x, but depends on α_1 and α_2, and $\alpha^{(2)}$ is identical with the corresponding function in (3) ($\nu = 2$).

We next give some simple applications of the theorems concerning the sums (4) and (5). Finally, we consider briefly the hitherto excluded case $\nu = 1$ of the sum (3): in place of an asymptotic formula we now get an identity. This identity suggests several others that are rather curious, and seem to be of some interest. It should be remarked that for special values of $\alpha(n)$, the identities obtained reduce to identities discussed by Evelyn and Linfoot (see § 6 for reference).

1. Definitions

Let $\alpha(n)$ be an arithmetic function which is

(i) factorable: that is,

$$\alpha(mn) = \alpha(m)\alpha(n) \text{ for } (m, n) = 1;$$

(ii) 'linear': that is,

$$\alpha(1) = 1, \qquad \alpha(n) = O(n^{\epsilon}) \quad (\epsilon > 0),$$
$$\alpha(p) = 1+O(p^{-\frac{1}{2}+\epsilon}),\dagger$$

where p, as throughout this paper, denotes a prime.

Let $\beta(n)$ be the so-called derivative of $\alpha(n)$:

$$\beta(n) = \sum_{n=d\delta} \mu(d)\alpha(\delta),$$

where, as usual, $\mu(d)$ is the Möbius function. It is clear that $\beta(n)$ has the following properties:

(i) $\beta(n)$ is factorable;

(iii)
$$\beta(1) = 1, \qquad \beta(n) = O(n^{\epsilon}),$$
$$\beta(p) = O(p^{-\frac{1}{2}+\epsilon}) \qquad (\epsilon > 0).$$

† The $\frac{1}{2}$ in the exponent might be replaced by any positive quantity; however, for simplicity, we limit ourselves to the case that gives the best error term in our theorems.

We shall say for brevity that any arithmetic function satisfying (i) and (ii) is an alpha-function; any function satisfying (i) and (iii) is a beta-function.

LEMMA. *If $\beta(m)$ is any beta-function, then*

$$\sum_{m\leqslant x}\beta(m) = O(x^{\frac{1}{2}+\epsilon}). \tag{7}$$

Define the function $\theta(m)$ thus:

$$\theta(1) = 1,$$
$$\theta(m) = 1 \text{ if for } p|m,\ p^2|m,$$
$$\theta(m) = 0 \text{ otherwise.}$$

Then we may verify that, for $s > \frac{1}{2}$,

$$\sum_{m=1}^{\infty}\frac{\theta(m)}{m^s} = \frac{\zeta(2s)\zeta(3s)}{\zeta(6s)} \qquad \left(\zeta(s) = \sum_{m=1}^{\infty}\frac{1}{m^s}\right),$$

whence it is easy to show that

$$\sum_{m\leqslant x}\theta(m) = O(x^{\frac{1}{3}}); \tag{8}$$

and indeed much more may be proved by elementary methods.

Now it is evident that any integer m may be written in the form

$$m = qh, \text{ where } \mu^2(q) = 1,\ \theta(h) = 1,\ (q,h) = 1.$$

Accordingly,

$$\begin{aligned}\left|\sum_{m\leqslant x}\beta(m)\right| &= \left|\sum_{\substack{qh\leqslant x\\(q,h)=1}}\mu^2(q)\theta(h)\beta(qh)\right| \\ &= \sum_{qh\leqslant x}\mu^2(q)|\beta(q)|\theta(h)|\beta(h)| \\ &= O\left\{\sum_{q\leqslant x}q^{-\frac{1}{2}+\epsilon}\sum_{h\leqslant x/q}\theta(h)h^{\epsilon}\right\} \\ &= O\left\{x^{2\epsilon}\sum_{q\leqslant x}q^{-\frac{1}{2}}(x/q)^{\frac{1}{3}}\right\}, \text{ by (8)}, \\ &= O\left\{x^{\frac{1}{3}+2\epsilon}\sum_{q\leqslant x}q^{-1}\right\} = O(x^{\frac{1}{3}+3\epsilon}).\end{aligned}$$

This proves the lemma.

Using (7), it is clear that the series

$$b(m) = \sum_{s=1}^{\infty}\frac{\beta(ms)}{s} \tag{9}$$

is absolutely convergent. We put

$$\tau(m) = \frac{b(m)}{b(1)}. \tag{10}$$

We shall now show that $\tau(m)$ is a beta-function. Evidently

$$b(m) = \prod_p \left\{\beta(p^e) + \frac{\beta(p^{e+1})}{p} + \dots\right\},$$

where $m = \prod p^e$, the product being taken over all p (so that $e = e(p) = 0$ for $p \nmid m$): it will be convenient to use this notation throughout. Then, by (10),

$$\tau(m) = \prod_{p|m} \frac{\beta(p^e) + \dfrac{\beta(p^{e+1})}{p} + \dots}{1 + \dfrac{\beta(p)}{p} + \dots}. \tag{11}$$

(We assume, of course, that $\beta(m)$ is such that no denominator in (11) vanishes; that is to say, $b(1) \neq 0$.) From (11) it is obvious that $\tau(m)$ is factorable. It is also fairly clear that $\tau(m)$ has the other necessary property, (iii) above. For the sequel, the relation between the beta-functions, $\beta(m)$, $\tau(m)$, is important, and we shall write, as an abbreviation for (11),

$$\tau = \Omega\beta. \tag{12}$$

As for the inverse $\beta = \Omega^{-1}\tau$, it is easy to see that, from (8) or (9), we get

$$\beta(m) = \frac{t(m)}{t(1)}, \text{ where } t(m) = \sum_{s=1}^{\infty} \frac{\mu(s)\tau(ms)}{s}, \tag{13}$$

so that

$$\beta(m) = \prod_{p|m} \frac{\tau(p^e) - \dfrac{\tau(p^{e+1})}{p}}{1 - \dfrac{\tau(p)}{p}}. \tag{14}$$

We remark that

$$b(1)t(1) = 1. \tag{15}$$

It is clear from the definition of a beta-function that the product of several beta's is itself a beta; also $m^{-1}\beta(m)$ is a beta whenever $\beta(m)$ is one; therefore, plainly

$$\tau^{(\nu)} = m^{1-\nu}\tau_1(m)\dots\tau_\nu(m) \tag{16}$$

is a beta, where, of course,

$$\tau_i = \Omega\beta_i \text{ for } i = 1,\dots,\nu.$$

Then, as we have seen above, the function $\beta^{(\nu)}$, defined by

$$\beta^{(\nu)} = \Omega^{-1}\tau^{(\nu)},$$

is a beta, and finally, therefore,

$$\alpha^{(\nu)}(m) = \sum_{d|m} \beta^{(\nu)}(m) \tag{17}$$

is an alpha-function.

If we write $b[\beta]$ for $b(1)$, defined by (9), and $t[\tau]$ for $t(1)$, defined by (13), or more briefly,

$$b_i = b[\beta_i], \qquad t^{(\nu)} = t[\tau^{(\nu)}], \tag{18}$$

then we may state our first principal theorem in the following form.

THEOREM 1.† *If $\alpha_1,\dots,\alpha_\nu$ be ν alpha-functions; $h_1,\dots,h_\nu$ be non-negative integers; then, for all $\nu \geqslant 2$, $\epsilon > 0$,*

$$\begin{aligned} R_\nu(n) &= \sum_{n=n_1+\dots+n_\nu} n_1^{h_1}\alpha_1(n_1)\dots n_\nu^{h_\nu}\alpha_\nu(n_\nu) \\ &= H_\nu b_1\dots b_\nu t^{(\nu)} n^{h_1+\dots+h_\nu+\nu-1}\alpha^{(\nu)}(n)+O(n^{h_1+\dots+h_\nu+\nu-\frac{1}{2}+\epsilon}), \end{aligned} \tag{19}$$

where $\alpha^{(\nu)}(n)$, b_i, $t^{(\nu)}$ are defined by (16), (17), *and* (18); *and*

$$H_\nu = \frac{(h_1+\dots+h_\nu+\nu-1)!}{h_1!\dots h_\nu!}.$$

We proceed now to prove the theorem first for $\nu = 2$, then, by an easy induction, for all $\nu \geqslant 2$.

2. Proof of Theorem 1 for $\nu = 2$.

2.1. From the definition of $\beta_i(m)$ we see that

$$R(n) = R_2(n) = \sum_{n=d_1e_1+d_2e_2} \beta_1(e_1)\beta_2(e_2)(d_1e_1)^{h_1}(d_2e_2)^{h_2}.$$

Let

$$R'(n) = \sum_{\substack{n=d_1e_1+d_2e_2 \\ e_1,e_2\leqslant n^{2/3}}} \beta_1(e_1)\beta_2(e_2)(d_1e_1)^{h_1}(d_2e_2)^{h_2}. \tag{20}$$

Then

$$|R(n)-R'(n)| < n^{h_1+h_2}\Big(\sum_{e_1>n^{2/3}} + \sum_{e_2>n^{2/3}}\Big)|\beta_1(e_1)\beta_2(e_2)|.$$

But

$$\begin{aligned} \sum_{\substack{n=d_1e_1+d_2e_2 \\ e_1>n^{2/3}}} |\beta_1(e_1)\beta_2(e_2)| &= \sum_{\substack{d_1e_1<n \\ e_1>n^{2/3}}} |\beta_1(e_1)| \sum_{d_2e_2=n-d_1e_1} |\beta_2(e_2)| \\ &= O(n^\epsilon)\sum_{e>n^{2/3}} |\beta_1(e)|\,\frac{n}{e} \end{aligned}$$

[for $\sum|\beta(d)|$, where $d|m$, is an alpha-function of m, and therefore $= O(m^\epsilon)$]

$$= O(n^{1+\epsilon})\sum_{e>n^{2/3}} \frac{|\beta_1(e)|}{e} = O(n^{\frac{1}{2}+\epsilon}),$$

by using (7). Therefore,

$$R(n)-R'(n) = O(n^{h_1+h_2+\frac{1}{2}+\epsilon}). \tag{21}$$

2.2. Now, from (20), we may write

$$\begin{aligned} R'(n) &= \sum_{m|n} \sum_{\substack{n=d_1e_1+d_2e_2 \\ (e_1,e_2)=m;\ e_1,e_2\leqslant n^{2/3}}} \beta_1(e_1)\beta_2(e_2)(d_1e_1)^{h_1}(d_2e_2)^{h_2} \\ &= \sum_{m|n} S'_m m^{h_1+h_2}, \text{ say,} \end{aligned} \tag{22}$$

† It is easily verified that, for $\alpha_1 = \dots = \alpha_\nu$, $\nu \geqslant 3$, $h_i = 0$, this theorem reduces to the result of I; the present error term is slightly better.

where
$$S'_m = \sum_{\substack{n/m=d_1f_1+d_2f_2\\(f_1,f_2)=1;\ f_1,f_2\leqq n^{2/3}m^{-1}}} (d_1f_1)^{h_1}(d_2f_2)^{h_2}\beta_1(mf_1)\beta_2(mf_2)$$
$$= \sum_{\substack{(f_1,f_2)=1\\f_1,f_2\leqq n^{2/3}m^{-1}}} f_1^{h_1}f_2^{h_2}\beta_1(mf_1)\beta_2(mf_2) \sum_{n/m=d_1f_1+d_2f_2} d_1^{h_1}d_2^{h_2}. \tag{23}$$

To evaluate the inner sum in (23), note that

$$d_1 \equiv d_1' \pmod{f_2} \qquad (0 \leqq d_1' < f_2),$$
$$d_2 \equiv d_2' \pmod{f_1} \qquad (0 \leqq d_2' < f_1),$$

so that the sum in question becomes

$$f_1^{h_1}f_2^{h_2} \sum_{n'=\delta_1+\delta_2} (\delta_1+\theta_1)^{h_1}(\delta_2+\theta_2)^{h_2}, \tag{24}$$

where
$$n' = \frac{n}{mf_1f_2} + O(1) \qquad (0 \leqq \theta_1,\ \theta_2 < 1).$$

But, by the Euler-Maclaurin sum-formula or otherwise, (24) reduces to

$$f_2^{h_1}f_1^{h_2}\left\{\frac{h_1!h_2!}{(h_1+h_2+1)!}\left(\frac{n}{mf_1f_2}\right)^{h_1+h_2+1} + O\left(\frac{n}{mf_1f_2}\right)^{h_1+h_2}\right\}$$
$$= H_2\left(\frac{n}{m}\right)^{h_1+h_2+1} f_1^{-h_1-1}f_2^{-h_2-1} + O\left\{\left(\frac{n}{m}\right)^{h_1+h_2} f_1^{-h_1}f_2^{-h_2}\right\}.$$

Substituting from the above, (23) becomes

$$S'_m = H_2\left(\frac{n}{m}\right)^{h_1+h_2+1} \sum_{\substack{(f_1,f_2)=1\\f_1,f_2\leqq n^{2/3}m^{-1}}} \frac{\beta_1(mf_1)\beta_2(mf_2)}{f_1f_2} +$$
$$+ O\left(\frac{n}{m}\right)^{h_1+h_2} \sum_{f_1,f_2\leqq n^{2/3}m^{-1}} |\beta_1(mf_1)\beta_2(mf_2)|.$$

But
$$\sum_{f\leqq n^{2/3}m^{-1}} |\beta(fm)| = O(mn^{\frac{2}{3}}m^{-1})^{\frac{1}{2}+\epsilon} = O(n^{\frac{1}{3}+\epsilon}).$$

Thus
$$S'_m = H_2\left(\frac{n}{m}\right)^{h_1+h_2+1} S''_m + O\left(\frac{n}{m}\right)^{h_1+h_2} n^{\frac{1}{3}+\epsilon}, \tag{25}$$

where
$$S''_m = \sum_{\substack{(f_1,f_2)=1\\f_1,f_2\leqq n^{2/3}m^{-1}}} \frac{\beta_1(mf_1)\beta_2(mf_2)}{f_1f_2}.$$

2.3. We now let

$$S_m = \sum_{\substack{f_1,f_2=1\\(f_1,f_2)=1}}^{\infty} \frac{\beta_1(mf_1)\beta_2(mf_2)}{f_1f_2};$$

then
$$|S_m - S''_m| < \sum_{f_1>n^{2/3}m^{-1}} \frac{|\beta_1(mf_1)|}{f_1} \sum_{f_2=1}^{\infty} \frac{|\beta_2(mf_2)|}{f_2} +$$
$$+ \sum_{f_1=1}^{\infty} \frac{|\beta_1(mf_1)|}{f_1} \sum_{f_2>n^{2/3}m^{-1}} \frac{|\beta_2(mf_2)|}{f_2}. \tag{26}$$

But $$\sum_1^\infty \frac{|\beta(mf)|}{f} = b\tau^*(m) = O(m^\epsilon) = O(n^\epsilon),$$

where $\tau^* = \Omega|\beta|$; on the other hand

$$\begin{aligned}\sum_{f>n^{\frac{1}{2}}m^{-1}} \frac{|\beta(mf)|}{f} &= O\sum_{f>n^{\frac{1}{2}}m^{-1}} \frac{(mf)^{\frac{1}{2}+\epsilon}}{f(f+1)} + O\left(\frac{n^{\frac{1}{2}+\epsilon}}{n^{\frac{1}{2}}m^{-1}}\right)\\ &= O\{m^{\frac{1}{2}+\epsilon}(n^{\frac{1}{2}}m^{-1})^{-\frac{1}{2}+\epsilon}\} + O(n^{-\frac{1}{2}+\epsilon}m^{1+\epsilon})\\ &= O(n^{-\frac{1}{2}+\epsilon}m^{1+\epsilon}).\end{aligned}$$

By (26), therefore,

$$S_m - S_m'' = O(n^{-\frac{1}{2}+\epsilon}m^{1+\epsilon}). \tag{27}$$

2.4. At this point we pause to combine the results of 2.1, 2.2, and 2.3. By (21),

$$\begin{aligned}R(n) &= R'(n) + O(n^{h_1+h_2+\frac{1}{2}+\epsilon})\\ &= \sum_{m|n} m^{h_1+h_2} S_m' + O(n^{h_1+h_2+\frac{1}{2}+\epsilon}), \text{ by (22)},\\ &= H_2 n^{h_1+h_2+1} \sum_{m|n} \frac{S_m''}{m} + O(n^{h_1+h_2+\frac{1}{2}+\epsilon}), \text{ by (25)},\\ &= H_2 n^{h_1+h_2+1} \sum_{m|n} \frac{S_m}{m} + O(n^{h_1+h_2+1} \sum_{m|n} n^{-\frac{1}{2}+\epsilon} m^\epsilon) +\\ &\qquad + O(n^{h_1+h_2+\frac{1}{2}+\epsilon}), \text{ by (27)},\\ &= H_2 n^{h_1+h_2+1} \sum_{m|n} \frac{S_m}{m} + O(n^{h_1+h_2+\frac{1}{2}+\epsilon}). \end{aligned} \tag{28}$$

Clearly, then, by the remarks made at the end of § 1—in particular, by (17)—in order to complete the proof of our theorem ($\nu = 2$), we have merely to show that

$$S_m = b_1 b_2 t^{(2)} m \beta^{(2)}(m). \tag{29}$$

This we now verify.

2.5. By the definition of S_m in 2.3, we have

$$\begin{aligned}\sum_{\delta=1}^\infty \frac{S_{m\delta}}{\delta^2} &= \sum_{\delta=1}^\infty \sum_{(f_1,f_2)=1} \frac{\beta_1(m\delta f_1)\beta_2(m\delta f_2)}{\delta^2 f_1 f_2}\\ &= \sum_{f_1,f_2=1}^\infty \frac{\beta_1(mf_1)\beta_2(mf_2)}{f_1 f_2}\\ &= b_1\tau_1(m) b_2\tau_2(m),\end{aligned} \tag{30}$$

as follows from formula (10). But inverting (30), we have immediately

$$S_m = b_1 b_2 \sum_{\delta=1}^\infty \frac{\mu(\delta)}{\delta^2} \tau_1(m\delta)\tau_2(m\delta)$$

$$= b_1 b_2 m \sum_{\delta=1}^{\infty} \frac{\mu(\delta)}{\delta} \tau^{(2)}(m\delta)$$

$$= b_1 b_2 t^{(2)} m \beta^{(2)}(m),$$

where

$$\tau^{(2)}(m) = \frac{\tau_1(m)\tau_2(m)}{m}, \qquad \beta^{(2)} = \Omega^{-1}\tau^{(2)}, \qquad t^{(2)} = t[\tau^{(2)}]$$

[see (16) and (18)]. We have then proved (29), thus completing the proof of the theorem in the case $\nu = 2$.

3. Proof of Theorem 1 for ν arbitrary

3.1. Let us assume now that Theorem 1 is true up to and including the value ν. Then we have

$$R_{\nu+1}(n) = \sum_{n=e+f} R_\nu(e)\alpha_{\nu+1}(f) f^{h_{\nu+1}}$$

$$= H_\nu b_1 \dots b_\nu t^{(\nu)} \sum_{n=e+f} e^{h_1+\dots+h_\nu+\nu-1} \alpha^{(\nu)}(e) f^{h_{\nu+1}} \alpha_{\nu+1}(f) + \\ + O(n^{h_1+\dots+h_{\nu+1}+\nu-\frac{1}{2}+\epsilon}), \quad (31)$$

the error term being estimated exactly as in the case of the sum in (24).

Since $\alpha^{(\nu)}(e)$ is an alpha-function, the sum in the right-hand member of (31) is an instance of (19) with $\nu = 2$. We write

$$h_1+\dots+h_\nu+\nu-1, \qquad h_{\nu+1}, \qquad \alpha^{(\nu)}, \qquad \alpha_{\nu+1},$$

in place of

$$h_1, \qquad h_2, \qquad \alpha_1, \qquad \alpha_2$$

respectively. Then the sum in question is

$$\sum_{n=e+f} = \frac{(h_1+\dots+h_\nu+\nu-1)!h_{\nu+1}!}{(h_1+\dots+h_{\nu+1}+\nu)!} b[\beta^{(\nu)}] b[\beta_{\nu+1}] t[\tau^*] \times \\ \times n^{h_1+\dots+h_{\nu+1}+\nu} \alpha^*(n) + O(n^{h_1+\dots+h_{\nu+1}+\nu-\frac{1}{2}+\epsilon}),$$

where

$$b[\beta^{(\nu)}] = \frac{1}{t[\tau^{(\nu)}]} = \frac{1}{t^{(\nu)}}, \quad \text{by (15);}$$

$$\tau^*(m) = \frac{\tau^{(\nu)}(m)\tau_{\nu+1}(m)}{m} = \frac{\tau_1(m)\dots\tau_\nu(m)\tau_{\nu+1}(m)}{m^\nu},$$

so that

$$\tau^* = \tau^{(\nu+1)}, \qquad t[\tau^*] = t^{(\nu+1)};$$

finally

$$\alpha^*(m) = \sum_{d|m} \beta^*(d), \qquad \beta^* = \Omega^{-1}\tau^*,$$

so that

$$\beta^* = \beta^{(\nu+1)} \quad \text{and} \quad \alpha^* = \alpha^{(\nu+1)}.$$

Substituting into (31), it is clear that the induction is complete.

3.2. We wish here to make two remarks about Theorem 1. In the first place, we note that the more general case in which the h_i are

merely real and each > -1 is by no means ruled out. The only places in which the proof is affected are those in which the Euler-Maclaurin formula is used. If, then, we make one obvious modification in the statement of the theorem, namely,

$$H_\nu = \frac{\Gamma(h_1+1)\ldots\Gamma(h_\nu+1)}{\Gamma(h_1+\ldots+h_\nu+\nu)},$$

we may still assert the truth of the theorem, at any rate with the less exact error term

$$O(n^{h_1+\ldots+h_\nu+\nu-1-c}) \text{ for some } c > 0.$$

Secondly, by modifying the proof in § 2 somewhat, it is not difficult to see that a direct proof of the case ν arbitrary may be given. Restricting ourselves, for simplicity, to the case in which all the h_i are zero, we see that the only difficulty is the evaluation of the inner sum in the general form of (23), or, what amounts to the same thing, the determination of the number of solutions in positive integers x_i of

$$n = a_1 x_1 + \ldots + a_\nu x_\nu,$$

where $(a_1,\ldots,a_\nu) = 1$; the number of solutions may be shown to be

$$N_\nu(n) = \frac{n^{\nu-1}}{(\nu-1)!a_1\ldots a_\nu} + O\left(\frac{n^{\nu-2}}{a_1\ldots a_k}\sum a_1 a_2\right).$$

I hope to show elsewhere in detail how this formula may be employed in a problem similar to that with which this paper is concerned.

4. The conjugate sum for $\nu = 2$

We now consider the sum (5)—in slightly more general form involving the h's.

THEOREM 2. *If k is a fixed positive integer; h_1, h_2 are non-negative integers; α_1, α_2 are arbitrary alpha-functions, then*

$$\begin{aligned}\bar{R}(x) &= \sum_{n\leqslant x} n^{h_1}\alpha_1(n)(n+k)^{h_2}\alpha_2(n+k)\\ &= \frac{x^{h_1+h_2+1}}{h_1+h_2+1} b_1 b_2 t^{(2)}\alpha^{(2)}(k) + O(x^{h_1+h_2+\frac{3}{4}+\epsilon}),\end{aligned}$$

where b_1, b_2, $t^{(2)}$, $\alpha^{(2)}$ are defined as in the case $\nu = 2$ of Theorem 1.

4.1. As in § 2.1, we begin with

$$\bar{R}(x) = \sum_{\substack{d_1e_1\leqslant x\\ d_1e_1+k=d_2e_2}} \beta_1(e_1)\beta_2(e_2)(d_1e_1)^{h_1}(d_2e_2)^{h_2}.$$

Let

$$R'(x) = \sum_{\substack{d_1e_1+k=d_2e_2\\ d_1e_1\leqslant x;\, e_1,e_2\leqslant x^{1/2}}} \beta_1(e_1)\beta_2(e_2)(d_1e_1)^{h_1}(d_2e_2)^{h_2}.$$

Then $$|\overline{R}(x)-R'(x)| < x^{h_1}(x+k)^{h_2}\Big(\sum_{e_1>x^{\frac{1}{2}}}+\sum_{e_2>x^{\frac{1}{2}}}\Big)|\beta_1(e_1)\beta_2(e_2)|.$$

But $$\sum_{e_1>x^{\frac{1}{2}}}|\beta_1(e_1)\beta_2(e_2)| = \sum_{\substack{d_1e_1\leqq x\\ e_1>x^{\frac{1}{2}}}}|\beta_1(e_1)|\sum_{d_1e_1+k=d_2e_2}|\beta_2(e_2)|$$

$$= O(x^{\epsilon})\sum_{e>x^{\frac{1}{2}}}|\beta_1(e)|\frac{x}{e} = O(x^{\frac{1}{2}+\epsilon}); \tag{32}$$

and similarly

$$\sum_{e_2>x^{\frac{1}{2}}}|\beta_1(e_1)\beta_2(e_2)| = \sum_{\substack{d_2e_2\leqq x+k\\ e_2>x^{\frac{1}{2}}}}|\beta_2(e_2)|\sum_{d_2e_2-k=d_1e_1}|\beta_1(e_1)| = O(x^{\frac{1}{2}+\epsilon}). \tag{33}$$

Therefore, by (32) and (33),

$$\overline{R}(x)-R'(x) = O(x^{h_1+h_2+\frac{1}{2}+\epsilon}). \tag{34}$$

4.2. We now write

$$R'(x) = \sum_{m|k} S'_m m^{h_1+h_2}, \tag{35}$$

where $$S'_m = \sum_{\substack{k/m=d_2f_2-d_1f_1\\ (f_1,f_2)=1\\ d_1f_1\leqq x;\ f_1,f_2\leqq x^{\frac{1}{2}}m^{-1}}}(d_1f_1)^{h_1}(d_2f_2)^{h_2}\beta_1(mf_1)\beta_2(mf_2).$$

Then proceeding exactly as at the corresponding point in § 2.2, we find that

$$S'_m = \frac{1}{h_1+h_2+1}\left(\frac{x}{m}\right)^{h_1+h_2+1}S''_m+O(x^{h_1+h_2+\frac{1}{2}+\epsilon}), \tag{36}$$

where $$S''_m = \sum_{\substack{(f_1,f_2)=1\\ f_1,f_2\leqq x^{\frac{1}{2}}m^{-1}}}\frac{\beta_1(mf_1)\beta_2(mf_2)}{f_1f_2}.$$

4.3. If we put $$S_m = \sum_{(f_1,f_2)=1}\frac{\beta_1(mf_1)\beta_2(mf_2)}{f_1f_2},$$

then, precisely as in 2.3,

$$S_m-S''_m = O(x^{-\frac{1}{2}+\epsilon}). \tag{37}$$

4.4. We now combine the results of 4.1, 4.2, 4.3.

$$\begin{aligned}\overline{R}(x) &= R'(x)+O(x^{h_1+h_2+\frac{1}{2}+\epsilon}) \quad \text{by (34)},\\ &= \sum_{m|k} m^{h_1+h_2}S'_m+O(x^{h_1+h_2+\frac{1}{2}+\epsilon}) \quad \text{by (35)},\\ &= \frac{x^{h_1+h_2+1}}{h_1+h_2+1}\sum_{m|k}\frac{S''_m}{m}+O(x^{h_1+h_2+\frac{1}{2}+\epsilon}) \quad \text{by (36)},\\ &= \frac{x^{h_1+h_2+1}}{h_1+h_2+1}\sum_{m|k}\frac{S_m}{m}+O(x^{h_1+h_2+\frac{1}{2}+\epsilon}) \quad \text{by (37)}.\end{aligned}$$

4.5. Evidently, in order to complete the proof of the theorem, it is only necessary to show that

$$S_m = b_1b_2t^{(2)}m\beta^{(2)}(m);$$

but this is simply (29), and the quantities involved in the two equations have exactly the same significance. Hence the result of 2.5 may be applied, and the proof is completed.

5. Applications

5.1. Let the arithmetic function $\mu_s(n)$, where $s(>1)$ is a fixed integer, be defined by

$$\mu_s(n) = 1 \quad \text{for} \quad p^s \nmid n,$$
$$\mu_s(n) = 0 \quad \text{for} \quad p^s \mid n;$$

it is easy to verify that $\mu_s(n)$ is an alpha-function. We apply the definitions of § 1 to calculate the corresponding β and τ.

(i) $$\beta(p^e) = -1 \quad \text{for} \quad e = s,$$
$$\beta(p^e) = 0 \quad \text{for} \quad 1 \leqslant e < s \text{ or } e > s.$$

(ii) $$\tau(p^e) = 0 \quad \text{for} \quad e > s,$$
$$\tau(p^e) = \frac{-p^e}{p^s - 1} \quad \text{for} \quad 1 \leqslant e \leqslant s.$$

Let now $s_1, \ldots, s_\nu$ be a set of ν integers each greater than 1; assume them so numbered that s_ν is not greater than any of the others. Defining the set of alpha-functions

$$\alpha_i(n) = \mu_{s_i}(n),$$

we shall now calculate $\tau^{(\nu)}$, $\beta^{(\nu)}$, and $\alpha^{(\nu)}$.

(iii) $$\tau^{(\nu)}(p^e) = 0 \quad \text{for} \quad e > s_\nu,$$
$$\tau^{(\nu)}(p^e) = \frac{(-1)^\nu p^e}{(p^{s_1}-1)\ldots(p^{s_\nu}-1)} \quad \text{for} \quad 1 \leqslant e \leqslant s_\nu.$$

(iv) $$\beta^{(\nu)}(p^e) = 0 \quad \text{for} \quad e \neq s_\nu \quad (e > 0),$$
$$\beta^{(\nu)}(p^e) = \frac{\dfrac{(-1)^\nu p^e}{(p^{s_1}-1)\ldots(p^{s_\nu}-1)}}{1 - \dfrac{(-1)^\nu}{(p^{s_1}-1)\ldots(p^{s_\nu}-1)}} \quad \text{for} \quad e = s_\nu.$$

(v) $$\alpha^{(\nu)}(p^e) = 0 \quad \text{for} \quad 1 \leqslant e < s_\nu,$$
$$\alpha^{(\nu)}(p^e) = \frac{1 + \dfrac{(-1)^\nu}{(p^{s_1}-1)\ldots(p^{s_{\nu-1}}-1)}}{1 - \dfrac{(-1)^\nu}{(p^{s_1}-1)\ldots(p^{s_\nu}-1)}} \quad \text{for} \quad e \geqslant s_\nu.$$

We may now state the theorem.

THEOREM 3.† *If $h_1,\dots, h_\nu$ are non-negative integers; $s_1,\dots, s_\nu$ integers each $\geqslant s_\nu > 1$; then the sum*

$$\sum_{n=n_1+\dots+n_\nu} n_1^{h_1}\dots n_\nu^{h_\nu}, \quad \text{where} \quad p^{s_i}\nmid n_i, \quad \nu \geqslant 2,$$

$$= H_\nu C\eta(n) n^{h_1+\dots+h_\nu+\nu-1} + O(n^{h_1+\dots+h_\nu+\nu-\frac{1}{2}+\epsilon}),$$

where

$$H_\nu = \frac{h_1!\dots h_\nu!}{(h_1+\dots+h_\nu+\nu-1)!},$$

$$C\zeta(s_1)\dots\zeta(s_\nu) = \prod_p \left\{1-\frac{(-1)^\nu}{(p^{s_1}-1)\dots(p^{s_\nu}-1)}\right\},$$

and

$$\eta(n) = \prod_{p^{s_\nu}\mid n}\left\{1-\frac{(-1)^{\nu-1}}{(p^{s_1}-1)\dots(p^{s_{\nu-1}}-1)}\right\}\left\{1-\frac{(-1)^\nu}{(p^{s_1}-1)\dots(p^{s_\nu}-1)}\right\}^{-1}.$$

5.2. We now consider a more general problem. Let $s(p)$ be a function of the prime p, whose values are integers not less than 1 always, and indeed not less than 2 except for at most a finite number of primes. Define the factorable function $g(n)$ by

$$g(1) = 1,$$
$$g(p^e) = 1 \quad \text{for} \quad e < s(p),$$
$$g(p^e) = 0 \quad \text{for} \quad e \geqslant s(p);$$

it is easy to see that $g(n)$ is an alpha-function. The corresponding β and τ are defined as follows:

(i) $\beta(p^e) = -1 \quad \text{for} \quad e = s(p),$

$\beta(p^e) = 0 \quad \text{for} \quad 1 \leqslant e < s(p) \text{ or } e > s(p);$

(ii) $\tau(p^e) = 0 \quad \text{for} \quad e > s(p),$

$$\tau(p^e) = \frac{-p^e}{p^{s(p)}-1} \quad \text{for} \quad 1 \leqslant e \leqslant s(p).$$

Let now $s_1(p),\dots, s_\nu(p)$ be a set of ν functions of the prime p, satisfying the conditions satisfied by $s(p)$ above. With each $s_i(p)$ we associate an alpha-function g_i: we have already calculated the corresponding β_i and τ_i; we shall now calculate $\tau^{(\nu)}$, $\beta^{(\nu)}$, $\alpha^{(\nu)}$.

† This theorem in the case $h_i = 0$, $s_1 = \dots = s_\nu$, is due to Evelyn and Linfoot, *Math. Zeitschrift*, 30 (1929), 433–48; see also I, p. 105. The case $\nu = 2$, $h = 0$, is due to A. Page, *Journal London Math. Soc.* 6 (1932): from this paper it will be clear how the error term in Theorem 3 may be improved.

(iii) $\tau^{(\nu)}(p^e) = 0 \quad \text{for} \quad e > s(p),$

$$\tau^{(\nu)}(p^e) = \frac{(-1)^\nu p^e}{(p^{s_1(p)}-1)\ldots(p^{s_\nu(p)}-1)} \quad \text{for} \quad 1 \leqslant e \leqslant s(p),$$

where
$$s(p) = \min_{1 \leqslant i \leqslant \nu} s_i(p); \tag{38}$$

(iv) $$\beta^{(\nu)}(p^e) = 0 \quad \text{for} \quad e \neq s(p), \quad e > 0,$$

$$\beta^{(\nu)}(p^e) = \frac{\dfrac{(-1)^\nu p^e}{(p^{s_1(p)}-1)\ldots(p^{s_\nu(p)}-1)}}{1-\dfrac{(-1)^\nu}{(p^{s_1(p)}-1)\ldots(p^{s_\nu(p)}-1)}} \quad \text{for} \quad e = s(p);$$

(v) $$\alpha^{(\nu)}(p^e) = 0 \quad \text{for} \quad 1 \leqslant e < s(p),$$

$$\alpha^{(\nu)}(p^e) = \frac{1+\dfrac{(-1)^\nu(p^{s(p)}-1)}{(p^{s_1(p)}-1)\ldots(p^{s_\nu(p)}-1)}}{1-\dfrac{(-1)^\nu}{(p^{s_1(p)}-1)\ldots(p^{s_\nu(p)}-1)}} \quad \text{for} \quad e \geqslant s(p).$$

THEOREM 4. *If $h_1,\ldots, h_\nu$ are non-negative integers; $s_1(p),\ldots, s_\nu(p)$ functions of the prime p such that (for each i) $s_i(p)$ is integral and $\geqslant 1$ for all p, and $s_i(p) \geqslant 2$ for all p with at most a finite number of exceptions; then the sum*

$$\sum_{n=n_1+\ldots+n_\nu} n_1^{h_1}\ldots n_\nu^{h_\nu}, \quad \text{where} \quad p^{s_i(p)} \nmid n_i, \ \nu \geqslant 2,$$

$$= H_\nu C\eta(n) n^{h_1+\ldots+h_\nu+\nu-1} + O(n^{h_1+\ldots+h_\nu+\nu-\frac{1}{2}+\epsilon}),$$

where H_ν is as in the previous theorem,

$$C = \prod_p \left\{1-\frac{(-1)^\nu}{(p^{s_1(p)}-1)\ldots(p^{s_\nu(p)}-1)}\right\} \prod_p (1-p^{-s_1(p)})\ldots(1-p^{-s_\nu(p)}),$$

and

$$\eta(n) = \prod_{p^{s_\nu(p)} \mid n} \left\{1-\frac{(-1)^{\nu-1}(p^{s(p)}-1)}{(p^{s_1(p)}-1)\ldots(p^{s_\nu(p)}-1)}\right\}\left\{1-\frac{(-1)^\nu}{(p^{s_1(p)}-1)\ldots(p^{s_\nu(p)}-1)}\right\}^{-1},$$

$s(p)$ being defined by (38).

5.3. Let $\sigma_s(n)$ denote the sum of the sth power of the divisors of n; evidently, for $s \leqslant -1/2$, σ_s is an alpha-function. For s merely negative, the corresponding function is one for which Theorem 1 still

applies, but with inferior error term. It is easy to verify that Theorem 1 implies the following result (see the first remark in § 3.2).

THEOREM 5. *Let $s_1,\dots, s_\nu$; $h_1,\dots, h_\nu$ denote (real) positive constants; then, for all $\nu \geqslant 2$,*

$$\sum_{n=n_1+\dots+n_\nu} n_1^{h_1-1}\sigma_{-s_1}(n_1)\dots n_\nu^{h_\nu-1}\sigma_{-s_\nu}(n_\nu)$$

$$\sim \frac{\zeta(s_1+1)\dots\zeta(s_\nu+1)}{\zeta(s+1)}\,\frac{\Gamma(h_1)\dots\Gamma(h_\nu)}{\Gamma(h_1+\dots+h_\nu)}\,n^{h_1+\dots+h_\nu-1}\sigma_{-s}(n),$$

where

$$s = s_1+\dots+s_\nu+\nu-1.$$

5.4. Theorem 5 is capable of a simple generalization: the generalization has the same relationship to the theorem that Theorem 4 has to Theorem 3.

Let $s(p)$ be a function of the prime p such that

$$s(p) \geqslant c > 0,\dagger$$

where c is independent of p. Then we define an alpha-function (in the more general sense: see § 5.3) by

$$\sigma'(p^e) = 1+p^{-s(p)}+\dots+p^{-es(p)}.$$

THEOREM 6. *Let $h_1,\dots, h_\nu$ denote positive constants; let $s_1(p),\dots, s_\nu(p)$ denote functions of the prime p such that*

$$s_i(p) \geqslant c > 0,\dagger$$

where c is an absolute constant; let $\sigma_i'(n)$ be a set of ν factorable functions of n defined by

$$\sigma_i'(p^e) = 1+p^{-s_i(p)}+\dots+p^{-es_i(p)};$$

then the sum

$$\sum_{n=n_1+\dots+n_\nu} n_1^{h_1-1}\dots n_\nu^{h_\nu-1}\sigma_1'(n_1)\dots\sigma_\nu'(n_\nu) \sim \frac{\Gamma(h_1)\dots\Gamma(h_\nu)}{\Gamma(h_1+\dots+h_\nu)}\,Bn^{h_1+\dots+h_\nu-1}\sigma'(n),$$

where

$$B = \prod_p \frac{1-p^{-s(p)-1}}{(1-p^{-s_1(p)-1})\dots(1-p^{-s_\nu(p)-1})},$$

$\sigma'(n)$ is factorable,

$$\sigma'(p^e) = 1+p^{-s(p)}+\dots+p^{-es(p)},$$

and

$$s(p) = s_1(p)+\dots+s_\nu(p)+\nu-1.$$

† It is sufficient to assume

$$s(p) \geqslant c\,\frac{\log\log p}{\log p}, \text{ where } c > 0.$$

5.5. If $\phi(n)$ is the Euler ϕ-function, then obviously $\phi(n)/n$ is an alpha-function. More generally, if $\phi_s(n)$ is Jordan's generalization of the ϕ-function:

$$\phi_s(n) = n^s \prod_{p|n} (1-p^{-s}),$$

then $n^{-s}\phi_s(n)$ is an alpha. Hence the following theorem:

THEOREM 7.† *If $h_1,\ldots, h_\nu$, $s_1,\ldots, s_\nu$ denote positive constants, then the sum*

$$\sum_{n=n_1+\ldots+n_\nu} n_1^{h_1-s_1-1}\ldots n_\nu^{h_\nu-s_\nu-1}\phi_{s_1}(n_1)\ldots\phi_{s_\nu}(n_\nu) \sim \frac{\Gamma(h_1)\ldots\Gamma(h_\nu)}{\Gamma(h_1+\ldots+h_\nu)} Bn^{h_1+\ldots+h_\nu-1}\eta(n),$$

where

$$B\zeta(s_1+1)\ldots\zeta(s_\nu+1) = \prod_p \left\{1-\frac{(-1)^\nu}{(p^{s_1+1}-1)\ldots(p^{s_\nu+1}-1)}\right\},$$

and $\eta(n)$ is an alpha-function defined by

$$\eta(n) = \prod_{p|n}\left\{1+\frac{(-1)^\nu(p-1)}{(p^{s_1+1}-1)\ldots(p^{s_\nu+1}-1)}\right\}\left\{1-\frac{(-1)^\nu}{(p^{s_1+1}-1)\ldots(p^{s_\nu+1}-1)}\right\}^{-1}.$$

It is clear how this theorem may be generalized by replacing each s by a function $s(p)$ that is greater than some positive constant: in view of the close analogy with previous theorems it is scarcely necessary to state the generalized theorem.

5.6. So far we have given applications of Theorem 1 only. However, comparison of Theorems 1 and 2 shows that for any application whatever of Theorem 1, the corresponding application of Theorem 2 may be written down immediately and without any further calculations. Therefore, we shall not take the space to enunciate the theorems conjugate to the case $\nu = 2$ of Theorems 3–7.

6. Identities arising from the case $\nu = 1$.

We have had no occasion in this paper to introduce the 'singular series' for the sum (2). It appears likely from the result of I that this should be

$$S_\nu(n) = \sum_{k=1}^{\infty} \frac{\tau^{(\nu)}(k)}{k} \sum_{\rho(k)} \rho^n = \sum_{k=1}^{\infty} \frac{\tau_1(k)\ldots\tau_\nu(k)}{k^\nu} \sum_{\rho(k)} \rho^n,$$

where $\rho = \rho(k)$ runs through the primitive kth roots of unity. Indeed, following I, § 4, it is easy to prove that

$$S_\nu(n) = t^{(\nu)}\alpha^{(\nu)}(n),$$

† For $\nu = 2$, see A. E. Ingham, *Journal London Math. Soc.* 2 (1927), 208.

so that the conclusion of Theorem 1 may be written (for $h_i = 0$) in the form

$$R_\nu(n) \sim \frac{b_1 \dots b_\nu}{(\nu-1)!} S_\nu(n) n^{\nu-1} \qquad (\nu \geqslant 2).$$

We remark that in the case $\nu = 1$ this asymptotic relation reduces to an identity. It is only necessary to show that

$$t^{(1)} = t_1 \quad \text{and} \quad \alpha^{(1)} = \alpha_1.$$

The first equation is obvious from the definition of $t^{(\nu)}$; as for the second, since $\tau^{(1)} = \tau_1$, it is evident that $\beta^{(1)} = \beta_1$, and therefore $\alpha^{(1)} = \alpha_1$. Dropping subscripts, we put this result in the form

$$\alpha(n) = bS(n) = b \sum_{k=1}^{\infty} \frac{\tau(k)}{k} \sum_{\rho(k)} \rho^n. \tag{39}$$

There is no difficulty in showing that the series converges.

Let now $f(x)$ be the generating function

$$f(x) = \sum_{k=1}^{\infty} \alpha(n) x^n \qquad (|x| < 1);$$

then, formally at least, from (39) we get

$$f(x) = bx \sum_{k=1}^{\infty} \frac{\tau(k)}{k} \sum_{\rho(k)} \frac{1}{\rho - x}. \tag{40}$$

This can be proved directly in the following way:

$$\begin{aligned} \sum_{n=1}^{\infty} \alpha(n) x^n &= \sum_{n=1}^{\infty} x^n \sum_{d \mid n} \beta(d) \\ &= \sum_{d=1}^{\infty} \frac{\beta(d) x^d}{1-x^d} \\ &= \sum_{d=1}^{\infty} \frac{x^d}{1-x^d} b \sum_{s=1}^{\infty} \frac{\mu(s)\tau(ds)}{s} \\ &= bx \sum_{k=1}^{\infty} \frac{\tau(k)}{k} \sum_{k=ds} \frac{dx^{d-1}\mu(s)}{1-x^d}. \end{aligned} \tag{41}$$

Let $F_k(x)$ denote the polynomial of lowest degree whose zeros are the primitive kth roots of unity; to make F_k unique, assume the term free of x to be $+1$; then, as is well known,

$$F_k(x) = \prod_{k=ds} (1-x^d)^{\mu(s)}.$$

It is obvious, then, that, using (41),

$$f(x) = -bx \sum_{k=1}^{\infty} \frac{\tau(k)}{k} \frac{F_k'(x)}{F_k(x)}, \tag{42}$$

which is equivalent to (40).

We next consider the sum

$$b_1 \sum_{k=1}^{\infty} \frac{\tau_1(k)}{k} \sum_{\rho(k)} f_2(x\rho), \quad \text{where} \quad f_i(x) = \sum_{n=1}^{\infty} \alpha_i(n) x^n.$$

By (39), this sum

$$= b_1 \sum_{k=1}^{\infty} \frac{\tau_1(k)}{k} \sum_{\rho(k)} \sum_{n=1}^{\infty} \alpha_2(n) x^n \rho^n$$

$$= \sum_{n=1}^{\infty} \alpha_2(n) x^n b_1 \sum_{k=1}^{\infty} \frac{\tau_1(k)}{k} \sum_{\rho(k)} \rho^n$$

$$= \sum_{n=1}^{\infty} \alpha_1(n) \alpha_2(n) x^n = \sum_{n=1}^{\infty} \alpha(n) x^n,$$

where $\alpha\,(= \alpha_1 \alpha_2)$ is itself an alpha-function. We have, therefore,

$$\sum_{n=1}^{\infty} \alpha_1(n) \alpha_2(n) x^n = b_1 \sum_{k=1}^{\infty} \frac{\tau_1(k)}{k} \sum_{\rho(k)} f_2(x\rho)$$

$$= b_2 \sum_{k=1}^{\infty} \frac{\tau_2(k)}{k} \sum_{\rho(k)} f_1(x\rho). \tag{43}$$

This result reduces to (40) on putting $\alpha_2(n) \equiv 1$.

As an illustration of these formulae, let $\alpha(n) = g(n)$, where $g(n)$ is the function defined in § 5.2. If we write $g\{n; s(p)\}$ in place of $g(n)$, then formulae (39), (40), (42), (43) reduce to the following direct generalizations of the Evelyn-Linfoot† identities:

$$g\{n; s(p)\} = b \sum_{k=1}^{\infty} g\{n; s(p)+1\} \prod_{p|k} \frac{-1}{p^{s(p)}-1} \sum_{\rho(k)} \rho^n; \tag{39'}$$

$$f(x) = \sum_{n=1}^{\infty} g\{n; s(p)\} x^n = bx \sum_{k=1}^{\infty} g\{k; s(p)+1\} \prod_{p|k} \frac{-1}{p^{s(p)}-1} \sum_{\rho(k)} \frac{1}{\rho - x} \tag{40'}$$

$$= -bx \sum_{k=1}^{\infty} g\{k; s(p)+1\} \prod_{p|k} \frac{-1}{p^{s(p)}-1} \frac{F_k(x)}{F_k'(x)}; \tag{42'}$$

† *Annals of Mathematics*, 32 (1931), 268–70. There are several other identities given by Evelyn and Linfoot that may be immediately generalized to the case of a general alpha-function.

and

$$\sum_{n=1}^{\infty} g\{n; s(p)\}x^n = b_1 \sum_{k=1}^{\infty} g\{k; s_1(p)+1\} \prod_{p|k} \frac{-1}{p^{s_1(p)}-1} \sum_{\rho(k)} f_2(x\rho)$$

$$= b_2 \sum_{k=1}^{\infty} g\{k; s_2(p)+1\} \prod_{p|k} \frac{-1}{p^{s_2(p)}-1} \sum_{\rho(k)} f_1(x\rho), \quad (43')$$

where $b = \prod_p (1-p^{-s(p)}), \qquad b_i = \prod_p (1-p^{-s_i(p)}),$

and in (43′) $s(p) = \min_{i=1,2} s_i(p).$

THE SECOND CURVATURE OF A SUB-SPACE

By A. G. WALKER (*Edinburgh*)

[Received 11 March 1932]

1. Introduction. This paper is mainly a discussion of the first and second curvatures of a Riemannian sub-space. The first curvature is well known as the mean curvature, and has been examined by Eisenhart and other writers; the second curvature is introduced directly as a generalization of the squared torsion of a twisted curve. The notation employed is, in general, that used by Eisenhart,* and we shall quote many of his results without giving explicit references.

2. Notation. Consider a flat space V_p, coordinates z^t $(t = 1, 2,..., p)$. Let V_m, coordinates y^α $(\alpha = 1, 2,..., m)$, be a sub-space of V_p, and let V_n, coordinates x^i $(i = 1, 2,..., n)$, be a sub-space of V_m. Then V_m, V_n are given by equations of the form

$$z^t = z^t(y); \quad y^\alpha = y^\alpha(x), \quad \text{or } z^t = z^t(x), \quad (t = 1, 2,..., p). \quad (1.1)$$

As V_p is a flat space, we can use the generalized vector analysis when referring to this space.† Let $\mathbf{r}$ be the position vector of a point of V_p. Then $\mathbf{r}$ is a function of the y's for points of V_m, and of the x's for points of V_n. We shall use the Greek letters α, β, γ, δ, and the roman letters h, i, j, k for suffixes when referring to V_m and V_n respectively. Thus we shall write $f_\alpha = \partial f/\partial y^\alpha$, and $f_i = \partial f/\partial x^i$. Also, $f_{,\alpha\beta}$ and $f_{,ij}$ will denote the second covariant derivatives of f with respect to V_m and V_n respectively.

The fundamental tensors of V_m, V_n are given by

$$a_{\alpha\beta} = \mathbf{r}_\alpha \cdot \mathbf{r}_\beta, \qquad g_{ij} = \mathbf{r}_i \cdot \mathbf{r}_j, \quad (1.2)$$

and from the relation $\mathbf{r}_i = \mathbf{r}_\alpha y^\alpha{}_{,i}$, we have at once $g_{ij} = a_{\alpha\beta} y^\alpha{}_{,i} y^\beta{}_{,j}$.

There are $m-n$ independent normals to V_n in V_m, and $p-m$ normals to V_m in V_p. We shall choose these to be mutually orthogonal, writing them $\mathbf{N}_{\sigma|}$ $(\sigma = 1, 2,..., m-n)$, and $\mathbf{N}_{r|}$ $(r = m-n+1, ..., p-n)$. Thus we have

$$\mathbf{N}_{\sigma|} \cdot \mathbf{N}_{\mu|} = 0, \quad \mathbf{N}^2_{\sigma|} = e_\sigma = \pm 1; \quad \mathbf{N}_{r|} \cdot \mathbf{N}_{s|} = 0, \quad \mathbf{N}^2_{r|} = e_r = \pm 1,$$

and

$$\mathbf{N}_{\sigma|} \cdot \mathbf{r}_i = 0, \quad \mathbf{N}_{r|} \cdot \mathbf{r}_\alpha = 0, \quad (1.3)$$

$$(\sigma, \mu = 1, 2,..., m-n, \quad r, s = m-n+1,..., p-n, \quad i = 1, 2,..., n,$$
$$\alpha = 1, 2,..., m, \quad \sigma \neq \mu; r \neq s).$$

* *Riemannian Geometry* (1926).

† For an introduction to this use of the vector analysis, see the writer's paper 'On deformation of sub-spaces': *Proc. Edinburgh Math. Soc.* (1932).

As the normals $\mathbf{N}_{\sigma|}$ are tangent to V_m, we can write

$$\mathbf{N}_{\sigma|} = \xi_{\sigma|}{}^{\alpha}\mathbf{r}_{\alpha} \qquad (\sigma = 1, 2, \ldots, m-n) \tag{1.31}$$

where $\xi_{\sigma|}{}^{\alpha}$ are the contravariant components of these normals in V_m.

The second fundamental tensors of V_n in V_m are given by

$$\left.\begin{aligned} \Omega_{\sigma|ij} &= \mathbf{N}_{\sigma|}\cdot\mathbf{r}_{,ij} = -\mathbf{N}_{\sigma|,i}\cdot\mathbf{r}_j = -\mathbf{N}_{\sigma|,j}\cdot\mathbf{r}_i \\ \mu_{\sigma\mu|i} &= \mathbf{N}_{\sigma|}\cdot\mathbf{N}_{\mu|,i} = -\mathbf{N}_{\sigma|,i}\cdot\mathbf{N}_{\mu|} \end{aligned}\right\}. \tag{1.4}$$

For V_n in V_p, we have the above tensors, together with the tensors $\Omega_{r|ij}$, $\mu_{rs|i}$ similarly defined, and also the tensors

$$\mu_{\sigma r|i} = \mathbf{N}_{\sigma|}\cdot\mathbf{N}_{r|,i}. \tag{1.41}$$

From (1.4) and (1.41) we see that the fundamental equations for V_n in V_p may now be written

$$\left.\begin{aligned} \mathbf{r}_{,ij} &= \sum_{\sigma} e_{\sigma}\Omega_{\sigma|ij}\mathbf{N}_{\sigma|} + \sum_{r} e_r\Omega_{r|ij}\mathbf{N}_{r|} \\ \mathbf{N}_{\sigma|,i} &= -\Omega_{\sigma|ij}g^{jk}\mathbf{r}_k - \sum_{\mu} e_{\mu}\mu_{\sigma\mu|i}\mathbf{N}_{\mu|} - \sum_{r} e_r\mu_{\sigma r|i}\mathbf{N}_{r|} \end{aligned}\right\}, \tag{1.5}$$

and the equations for V_n in V_m are found from these by omitting vector components normal to V_m.

For V_n in V_p, we have introduced the vector operator $\nabla = g^{ij}\mathbf{r}_i\dfrac{\partial}{\partial x^j}$. We can now extend this for V_n in V_m, and define the *relative operator*

$$\nabla = g^{ij}\mathbf{r}_i\frac{\partial_m}{\partial x^j},$$

where $\partial_m/\partial x^j$ is $\partial/\partial x^j$ followed by an operator for annihilating vector components normal to V_m.* In this way, we shall extend the operator $\nabla^2 = \nabla\cdot\nabla$, etc.

LEMMA. *To show that*

$$g^{ij}y^{\alpha}{}_{,i}y^{\beta}{}_{,j} = a^{\alpha\beta} - \sum_{\sigma} e_{\sigma}\xi_{\sigma|}{}^{\alpha}\xi_{\sigma|}{}^{\beta}.$$

Let $\lambda_{h|}{}^{i}$ $(i, h = 1, 2, \ldots, n)$ be the components of an orthogonal n-uple in V_n. Then the components in V_m are $\lambda_{h|}{}^{\alpha} = \lambda_{h|}{}^{i}y^{\alpha}{}_{,i}$, and these n vectors, together with the $(m-n)$ vectors $\xi_{\sigma|}{}^{\alpha}$, form an orthogonal m-uple in V_m. Hence, we have

$$a^{\alpha\beta} = \sum_{h} e_h\lambda_{h|}{}^{\alpha}\lambda_{h|}{}^{\beta} + \sum_{\sigma} e_{\sigma}\xi_{\sigma|}{}^{\alpha}\xi_{\sigma|}{}^{\beta}. \tag{2.1}$$

But

$$\sum_{h} e_h\lambda_{h|}{}^{\alpha}\lambda_{h|}{}^{\beta} = y^{\alpha}{}_{,i}y^{\beta}{}_{,j}\sum_{h} e_h\lambda_{h|}{}^{i}\lambda_{h|}{}^{j} = g^{ij}y^{\alpha}{}_{,i}y^{\beta}{}_{,j}. \tag{2.11}$$

* For a discussion of the vector operator, see the writer's paper above mentioned. We shall use the same notation ∇ for the relative operator: for this is the only one with which we shall be concerned.

Therefore $$g^{ij}y^{\alpha}{}_{,i}y^{\beta}{}_{,j} = a^{\alpha\beta} - \sum_{\sigma} e_{\sigma}\xi_{\sigma|}{}^{\alpha}\xi_{\sigma|}{}^{\beta}. \tag{2.2}$$

If $m = n+1$, this becomes

$$\xi^{\alpha}\xi^{\beta} = e(a^{\alpha\beta} - g^{ij}y^{\alpha}{}_{,i}y^{\beta}{}_{,j}) = c^{\alpha\beta}, \tag{2.3}$$

say. Hence the components of the normal to V_n in V_{n+1} are given by

$$\xi^{\alpha} = e_{\alpha}(c^{\alpha\alpha})^{1/2}, \tag{2.4}$$

where $e_1 = 1$, and e_{α} is the sign of $c^{1\alpha}$.

For a surface in V_3, (2.4) becomes, with the usual notation,

$$X = \{1-(Gx_1^2 - 2Fx_1x_2 + Ex_2^2)\}^{1/2}, \text{ etc.}, \tag{2.5}$$

where $x_1 = \partial x/\partial u$, $x_2 = \partial x/\partial v$.

3. For a curve in V_3, the vector operator reduces to $\nabla = \mathbf{t}d/ds$, where $\mathbf{t}$ is the unit-tangent vector. We find that the curvature vector of the curve can be written

$$\kappa\mathbf{n} = \nabla^2\mathbf{r},$$

where $\mathbf{n}$ is the principal normal, and κ the curvature. Generalizing this expression to V_n in V_m, we write

$$M\mathbf{N} = \nabla^2\mathbf{r} = g^{ij}\sum_{\sigma} e_{\sigma}\Omega_{\sigma|ij}\mathbf{N}_{\sigma|}. \tag{3.1}$$

We see that $\mathbf{N}$ so defined is the mean-curvature normal, and M is the mean curvature of V_n in V_m, these, then, being the generalizations of the principal normal and the curvature of a curve. We shall call $\mathbf{N}$ the *principal normal*, and M the *first curvature* of V_n in V_m. The second fundamental tensor for the normal $\mathbf{N}$ we shall write as Ω_{ij}. Hence,

$$M\Omega_{ij} = \sum_{\sigma} e_{\sigma}\Omega_{\sigma|}\Omega_{\sigma|ij}; \qquad \Omega_{\sigma|} = g^{ij}\Omega_{\sigma|ij}, \tag{3.11}$$

and

$$M = e\Omega, \qquad M^2 = e\sum_{\sigma} e_{\sigma}\Omega_{\sigma|}{}^2; \qquad \Omega = g^{ij}\Omega_{ij}, \qquad \mathbf{N}^2 = e = \pm 1.$$

Also, for a curve, we find that the torsion is given by

$$\tau^2 = -\mathbf{n}\cdot\nabla^2\mathbf{n} - (\nabla\cdot\mathbf{n})^2,$$

where $\mathbf{n}$ is the principal normal. Hence, generalizing, we shall define the *second curvature* T of V_n in V_m by

$$T = -\mathbf{N}\cdot\nabla^2\mathbf{N} - (\nabla\cdot\mathbf{N})^2, \tag{3.2}$$

where $\mathbf{N}$ is the principal normal of V_n in V_m. We have*

$$\mathbf{N}\cdot\nabla^2\mathbf{N} = g^{ij}\mathbf{N}\cdot[\mathbf{N}_{,ij}]_m = -g^{ij}[\mathbf{N}_{,i}\cdot\mathbf{N}_{,j}]_m \tag{3.21}$$

* The suffix m denotes that, after each operation on the vectors in the brackets, the components normal to V_m must be omitted. This follows from the definition of the relative operators.

and $$(\nabla \cdot \mathbf{N})^2 = (g^{ij}\mathbf{r}_i \cdot \mathbf{N}_j)^2 = M^2.$$

Hence, choosing the remaining $m-n-1$ normals $\mathbf{N}_{\sigma|}$ to be orthogonal to $\mathbf{N}$, and writing $\mu_{\sigma|i} = \mathbf{N} \cdot \mathbf{N}_{\sigma|,i}$, we have, from (3.5),

$$T+M^2 = g^{ih}g^{jk}\Omega_{ij}\Omega_{hk}+g^{ij}\sum_{\sigma}\mu_{\sigma|i}\mu_{\sigma|j},$$

and hence,

$$T = g^{ih}g^{jk}(\Omega_{ij}\Omega_{hk}-\Omega_{ih}\Omega_{jk})+g^{ij}\sum_{\sigma}\mu_{\sigma|i}\mu_{\sigma|j}. \tag{3.22}$$

From (3.21), we see that an alternative definition would be

$$T+M^2 = S(\nabla\mathbf{N}\cdot\mathbf{N}\nabla), \tag{3.23}$$

where $\mathbf{N}\nabla$ is the dyadic conjugate to $\nabla\mathbf{N}$, and S denotes the scalar.

If $m = n+1$, we have

$$T = g^{ih}g^{jk}(\Omega_{ij}\Omega_{hk}-\Omega_{ih}\Omega_{jk}). \tag{3.3}$$

The Gauss equation for V_n in V_{n+1} is

$$e(\Omega_{ij}\Omega_{hk}-\Omega_{ih}\Omega_{jk}) = R_{ikjh}-R'_{\alpha\beta\gamma\delta}y^{\alpha}{}_{,i}y^{\beta}{}_{,k}y^{\gamma}{}_{,j}y^{\delta}{}_{,h}, \tag{3.31}$$

where R, R' are the curvature tensors of V_n, V_{n+1} respectively. Hence, (3.3) becomes

$$T = eR-eg^{ih}g^{jk}y^{\alpha}{}_{,i}y^{\beta}{}_{,k}y^{\gamma}{}_{,j}y^{\delta}{}_{,h}R'_{\alpha\beta\gamma\delta},$$

i.e. by (2.3), $$T = eR-e(a^{\alpha\delta}-e\xi^{\alpha}\xi^{\delta})(a^{\beta\gamma}-e\xi^{\beta}\xi^{\gamma})R'_{\alpha\beta\gamma\delta},$$

i.e. $$T = eR-eR'+2R'_{\alpha\beta}\xi^{\alpha}\xi^{\beta}, \tag{3.32}$$

where $R'_{\alpha\beta}$ is the Ricci tensor for V_m, and R, R' the scalar curvatures of V_n, V_m.

If V_{n+1} is an Einstein space, we have $R'_{\alpha\beta} = \lambda a_{\alpha\beta}$, and (3.32) reduces to

$$T = eR-e(n-1)\lambda. \tag{3.33}$$

If the Ricci tensor vanishes in V_{n+1}, we have

$$T = eR. \tag{3.34}$$

In particular, this is satisfied if V_{n+1} is a flat space. For the case of a surface in V_3, we have

$$T = R = -2K, \tag{3.35}$$

where K is the total curvature. Weatherburn* showed that $-2K$ was a generalization of the squared torsion, and the mean curvature a generalization of the curvature of a curve.

We observe that, for a curve in hyper-space, T is the squared second curvature, and, from (3.33), this vanishes for a curve in V_2, the curvature tensor of a curve being zero.

* *Messenger of Math.* 56 (1927), 173.

4. If the second fundamental tensors of V_n, V_m in V_p for the normals $\mathbf{N}_{r|}$ are $\Omega_{r|ij}$, $\Omega_{r|\alpha\beta}$, we have

$$\Omega_{r|\alpha\beta} = -\mathbf{N}_{r|,\alpha} \cdot \mathbf{r}_\beta$$

and
$$\Omega_{r|ij} = -\mathbf{N}_{r|,i} \cdot \mathbf{r}_j = -\mathbf{N}_{r|,\alpha} \cdot \mathbf{r}_\beta y^\alpha{}_{,i} y^\beta{}_{,j}.$$

Hence
$$\Omega_{r|ij} = \Omega_{r|\alpha\beta} y^\alpha{}_{,i} y^\beta{}_{,j}. \tag{4.1}$$

Also, from (1.41), we have

$$\mu_{\sigma r|i} = \mathbf{N}_{\sigma|} \cdot \mathbf{N}_{r|,i} = \xi_{\sigma|}{}^\alpha \mathbf{r}_\alpha \cdot \mathbf{N}_{r|,\beta} y^\beta{}_{,i},$$

i.e.
$$\mu_{\sigma r|i} = -\Omega_{r|\alpha\beta} \xi_{\sigma|}{}^\alpha y^\beta{}_{,i}. \tag{4.11}$$

If $\mathbf{N}$, $\mathbf{N}'$ are the principal normals and M, M' the first curvatures of V_n in V_m and V_n in V_p respectively, we have

$$M\mathbf{N} = \sum_\sigma e_\sigma \Omega_{\sigma|} \mathbf{N}_{\sigma|},$$

and
$$M'\mathbf{N}' = \sum_\sigma e_\sigma \Omega_{\sigma|} \mathbf{N}_{\sigma|} + \sum_r e_r \Omega_{r|} \mathbf{N}_{r|}. \tag{4.2}$$

Hence, the principal normal of V_n in V_p is tangent to V_m if

$$\Omega_{r|} = 0 \qquad (r = m-n+1, \ldots, p-n). \tag{4.21}$$

In this case, we see that

$$\mathbf{N}' = \mathbf{N}, \qquad M' = M, \qquad \Omega'_{ij} = \Omega_{ij}, \tag{4.22}$$

and we may say that V_n is an *asymptotic sub-space* of V_m. These results are evidently true if V_p is not necessarily a flat space.

Writing $\mu_{r|i} = \mathbf{N} \cdot \mathbf{N}_{r|,i}$, from (3.22), (4.22) we see that, if T, T' are the second curvatures of V_n in V_m, V_n in V_p, we have

$$T' - T = g^{ij} \sum_r e_r \mu_{r|i} \mu_{r|j}, \tag{4.23}$$

i.e. from (4.11),

$$T' - T = g^{ij} y^\alpha{}_{,i} y^\beta{}_{,j} \xi^\gamma \xi^\delta \sum_r e_r \Omega_{r|\alpha\gamma} \Omega_{r|\beta\delta}, \tag{4.24}$$

where $\mathbf{N} = \xi^\alpha \mathbf{r}_\alpha$. From (4.1), the equations (4.21) can be written

$$g^{ij} y^\alpha{}_{,i} y^\beta{}_{,j} \Omega_{r|\alpha\beta} = 0.$$

Hence, (4.24) can be written

$$T' - T = g^{ij} y^\alpha{}_{,i} y^\beta{}_{,j} \xi^\gamma \xi^\delta \sum_r e_r (\Omega_{r|\alpha\gamma} \Omega_{r|\beta\delta} - \Omega_{r|\alpha\beta} \Omega_{r|\gamma\delta}). \tag{4.25}$$

This equation is true if V_p is not a flat space. If V_p is a flat space, the Gauss equations for V_m in V_p are

$$\sum_r e_r (\Omega_{r|\alpha\gamma} \Omega_{r|\beta\delta} - \Omega_{r|\alpha\beta} \Omega_{r|\gamma\delta}) = R'_{\alpha\delta\gamma\beta}. \tag{4.26}$$

Hence, from (2.2), (4.25) becomes

$$T' - T = R'_{\alpha\beta} \xi^\alpha \xi^\beta - R'_{\alpha\delta\gamma\beta} \xi^\gamma \xi^\delta \sum_\sigma e_\sigma \xi_{\sigma|}{}^\alpha \xi_{\sigma|}{}^\beta. \tag{4.27}$$

If $m = n+1$, we have $\xi_1{}^\alpha = \xi^\alpha$, and (4.27) reduces to

$$T' - T = R'_{\alpha\beta}\xi^\alpha\xi^\beta. \tag{4.28}$$

Hence, from (3.32), we have

$$T' = eR - eR' + 3R'_{\alpha\beta}\xi^\alpha\xi^\beta. \tag{4.29}$$

We observe that the direction of the vector ξ^α is the principal normal direction of V_n in V_p, and T' is the second curvature of V_n in V_p. Also, R depends only on V_n. Hence

Given a space V_n in a flat space V_p, then all the spaces V_{n+1} which contain V_n and the principal normal direction of V_n in V_p at points of V_n have the same value for the expression $3R'_{\alpha\beta}\xi^\alpha\xi^\beta - eR'$ at points of V_n, where $R'_{\alpha\beta}$, R' refer to V_{n+1}, and ξ^α are the components in V_{n+1} of the principal normal of V_n in V_p.

This could be written

All spaces V_{n+1} of which V_n is an asymptotic sub-space have the same value for the above expression at points of V_n.

For a curve on a surface in V_3, we have $n = 1$, $m = 2$, $p = 3$, and $R'_{\alpha\beta} = -Ka_{\alpha\beta}$ where K is the total curvature. Hence we have

$$T' = -K,$$

R being zero for a curve. This is a well-known relation, for T' is the squared torsion of the curve, and the curve is an asymptotic line on the surface.

If V_p is any space, not necessarily flat, and if V_m is a sub-space of V_p, V_m is said to be *totally geodesic* if all geodesics in V_m are geodesics in V_p. The necessary and sufficient conditions for this are easily shown to be

$$\Omega_{r|\alpha\beta} = 0 \qquad \begin{pmatrix} r = 1, 2, \ldots, p-m \text{ or } = m-n+1, \ldots, p-n, \\ \alpha, \beta = 1, 2, \ldots, m. \end{pmatrix} \tag{4.3}$$

From (4.26), we see that if V_p is a flat space, then all totally geodesic sub-spaces are flat spaces.

If V_m is a totally geodesic sub-space of V_p, then (4.24) becomes

$$T' - T = 0, \tag{4.31}$$

i.e. *if V_m is a geodesic sub-space of a space V_p, and if V_n is immersed in V_m, then the second curvature of V_n in V_p is the second curvature of V_n in V_m.*

For the case $n = 1$, $m = 2$, $p = 3$, we have $T = 0$ and hence

$T' = 0$. If V_3 is Euclidean, V_2 is a plane, and the above theorem reduces to the property that the torsion of a plane curve is zero.

5. Consider a system of ∞^{m-n} spaces V_n in V_m, such that one passes through each point of V_m. Then the normals $\mathbf{N}_{\sigma|}$ may be considered as functions of position in V_m. We have $\mathbf{r}_i = \mathbf{r}_\alpha y^\alpha{}_{,i}$, $\mathbf{N}_{\sigma|,j} = \mathbf{N}_{\sigma|,\beta} y^\beta{}_{,j}$, and hence, for the normal $\mathbf{N}_{\sigma|} = \xi_{\sigma|}{}^\alpha \mathbf{r}_\alpha$, we have

$$\Omega_{\sigma|} = -g^{ij}\mathbf{r}_i \cdot \mathbf{N}_{\sigma|,j} = -g^{ij} y^\alpha{}_{,i} y^\beta{}_{,j}\, \mathbf{r}_\alpha \cdot \mathbf{N}_{\sigma|,\beta}, \tag{5.1}$$

i.e. substituting for $\mathbf{N}_{\sigma|}$,

$$\Omega_{\sigma|} = -g^{ij} y^\alpha{}_{,i} y^\beta{}_{,j} \xi_{\sigma|\alpha,\beta}. \tag{5.11}$$

We have $$\xi_{\mu|}{}^\alpha \xi_{\mu|}{}^\beta \xi_{\sigma|\alpha,\beta} = -\xi_{\mu|}{}^\alpha{}_{,\beta} \xi_{\mu|}{}^\beta \xi_{\sigma|\alpha} = -\eta_{\mu|}{}^\alpha \xi_{\sigma|\alpha},$$

where $\eta_{\mu|}{}^\alpha = \xi_{\mu|}{}^\alpha{}_{,\beta} \xi_{\mu|}{}^\beta$ is the principal or geodesic curvature vector of the system of curves $\xi_{\mu|}{}^\alpha$ in V_m. Hence, from (2.2), we have

$$\Omega_{\sigma|} = -\xi_{\sigma|}{}^\alpha{}_{,\alpha} - \sum_\mu e_\mu a_{\alpha\beta} \xi_{\sigma|}{}^\alpha \eta_{\mu|}{}^\beta. \tag{5.12}$$

We have thus found the mean curvatures for the normals as functions of position of V_m. Obtaining the $m-n$ mean curvatures in this way, we can now find the principal normal and the first curvature from the relations

$$M\xi^\alpha = \sum e_\sigma \Omega_{\sigma|} \xi_{\sigma|}{}^\alpha, \qquad M^2 = e \sum e_\sigma \Omega_{\sigma|}{}^2. \tag{5.13}$$

Having found the principal normal and first curvature, we can find the second curvature from the equation

$$T + M^2 = g^{ij}[\mathbf{N}_{,i} \cdot \mathbf{N}_{,j}]_m = g^{ij} y^\alpha{}_{,i} y^\beta{}_{,j} [\mathbf{N}_{,\alpha} \cdot \mathbf{N}_{,\beta}]_m, \tag{5.2}$$

i.e., substituting from $\mathbf{N} = \xi^\alpha \mathbf{r}_\alpha$, we have

$$T = -M^2 + \Big(a^{\alpha\beta} - \sum_\mu e_\mu \xi_{\mu|}{}^\alpha \xi_{\mu|}{}^\beta\Big) a_{\gamma\delta}\, \xi^\gamma{}_{,\alpha}\, \xi^\delta{}_{,\beta}. \tag{5.21}$$

If $m = n+1$, there is one normal ξ^α, this being the principal normal, and we have

$$\Omega = -\xi^\alpha{}_{,\alpha}, \qquad M = e\Omega, \tag{5.3}$$

and

$$T = -M^2 + (a^{\alpha\beta} - e\xi^\alpha \xi^\beta) a_{\gamma\delta}\, \xi^\gamma{}_{,\alpha}\, \xi^\delta{}_{,\beta}, \tag{5.31}$$

i.e.

$$T = -M^2 + a^{\alpha\beta} a_{\gamma\delta}\, \xi^\gamma{}_{,\alpha}\, \xi^\delta{}_{,\beta} - ee'\kappa^2, \tag{5.32}$$

where κ is the principal curvature of the curves of congruence ξ^α, and $e'\kappa^2 = a_{\gamma\delta}\eta^\gamma\eta^\delta$, $\eta^\alpha = \xi^\alpha{}_{,\beta}\xi^\beta$.

From the above equations, we see that, given $m-n$ directions, which we may take to be mutually orthogonal, at each point of V_m, then, if these satisfy the conditions of integrability of the system of equations $a_{\alpha\beta}\xi_{\sigma|}{}^\alpha dy^\beta = 0$ $(\sigma = 1, 2, \ldots, m-n)$, we can find the

principal normal and the curvatures of the spaces V_n defined as being normal to these $m-n$ directions.

Consider the case $m = 3$, $n = 2$, V_3 being referred to Cartesian axes. Let the given direction be (X, Y, Z), the components being functions of x, y, z.

The condition of integrability of $\sum X\,dx = 0$ is

$$X(Y_3-Z_2)+Y(Z_1-X_3)+Z(X_2-Y_1) = 0 \tag{5.4}$$

where $X_1 = \partial X/\partial x$, etc. If this is satisfied, the above directions are normal to a family of surfaces, and, from (5.3), (5.31), we see that the mean curvature K_m, and the total curvature, K, of these surfaces are given by

$$K_m = -(X_1+Y_2+Z_3), \tag{5.41}$$

$$2K = (X_1+Y_2+Z_3)^2 + \sum[(XX_1+YX_2+ZX_3)^2-(X_1^2+X_2^2+X_3^2)]. \tag{5.42}$$

As an example, consider

$$X = \frac{x}{r}, \qquad Y = \frac{y}{r}, \qquad Z = \frac{z}{r}, \qquad r^2 = x^2+y^2+z^2.$$

In this case, (5.41) reduces to $K_m = -2/r$, and (5.42) reduces to $K = 1/r^2$. These are the results we should expect, for the given directions are evidently normal to a family of concentric spheres.

THE THEORY OF THE OSCILLATIONS OF CERTAIN STELLAR MODELS ABOUT THEIR STEADY-STATE CONFIGURATIONS (I)

By L. E. LEFEVRE (*Pasadena*)

[Received 7 March 1932]

1. Introduction. The small oscillations about an equilibrium configuration of stellar models constructed according to certain hypotheses as to physical structure, generation of internal energy, etc., have been considered by Eddington,* Jeans,† L. H. Thomas,‡ and others. The investigations of Eddington apply a theory of 'adiabatic' pulsations of a particular model to the problem of the Cepheid variables. The oscillations of the interior regions of the star are supposed to proceed according to the 'adiabatic' law, $p\propto\rho^{\gamma}$, and, subsequently, an estimation is made of the error involved in carrying this assumption into the outer layers, by calculating the heat 'leakage' there. A more satisfactory procedure is to make use of an equation of energy-exchanges which accurately expresses the balance at any point between the net gain of energy per unit mass due to generation of internal energy, less loss by radiation, and the various transformations of the different kinds of material energy by physical processes. Such an equation was used by Jeans,§ who obtained a third-order differential equation with respect to the time for the oscillations of certain models, and derived therefrom certain stability criteria.

The outstanding difficulty of all existing theories of Cepheid variation is their failure to account for the observed relation between the phases of the variations in displacement and luminosity. According to Eddington's theory, the star throughout should be hottest when most compressed, whereas observational evidence shows that at the

* *M.N.* 79 (1918), 2, 177.

† *M.N.* 85 (1925), 914; 87 (1927), 400, 720.

‡ *M.N.* 91 (1930), 122, and 619 (1931). These papers have appeared since the present investigation was begun. The fundamental equations of motion used by Thomas are similar to, but not identical with, those used here (see footnote, p. 305). The development of the argument differs widely from the present one, however, being directed toward a discussion of the nature of the stability criteria for stellar configurations.

§ e.g. *Astronomy and Cosmogony* (1928), Chap. IV.

surface the temperature is greatest when the velocity of approach of the surface towards the observer is greatest; that is to say, the luminosity there is in phase with dr/dt instead of r.

In the general theory of small oscillations of systems possessing an infinite number of degrees of freedom, such as the vibrations of taut strings, organ-pipes, etc., the solution is obtained in terms of normal modes of oscillation, the most general motion of the system being a linear combination of such normal modes, whilst the fundamental periods of the system are the roots of a certain equation which has to be determined from the boundary conditions of the problem. The boundary conditions in these problems are always such that they determine just one such equation, the 'period equation' of the system. Similarly, in atomic theory, the Wave Mechanics requires a study of the 'eigenwerte' and 'eigenfunktionen' of the differential equations. A rational theory of stellar oscillations, capable of application to variable stars, should aim at a differential equation, or system of differential equations, whose solution can be expressed in terms of normal modes of oscillation. No existing theory has established the existence of normal modes of oscillation of the radial pulsations of a spherical mass of gas evolving energy, nor has any placed the emphasis upon the influence of boundary conditions in determining the period or periods of such a system that is suggested by the analogy of the examples quoted.

It is the object of the present paper to prove that, in accordance with general dynamical theory, there exists a definite 'period equation' for certain particular stellar models. This is proved for configurations built on the 'standard model', both in the case of the completely gaseous (Eddington) model and also for certain 'cored' (Milne) models.* The existence of this equation is derived from a complete consideration of the 'boundary' conditions in each case. As, however, the differential equations of the motion are not soluble explicitly, the period equation is not given formal mathematical expression. It further appears from these considerations that a definite relationship exists throughout the model between the phases of the variations in displacement and temperature. The determination of this relationship would require the very troublesome numerical quadrature of the differential equations of the motion for certain

* These models are chosen by way of illustration of the method only, and need not necessarily represent stars actually occurring in Nature.

suitable values of the disposable constants of the model. This has not yet been undertaken.

The method adopted is similar to that of Jeans,* but the treatment differs from his in that we replace his single equation by a system of equations of lower order.

In § 2 we briefly recapitulate for purposes of reference the equations of equilibrium of the gaseous layers of the Standard Model configuration.† The equations of motion for small oscillations of these gaseous layers are obtained in § 3. In § 4 we derive the system of equations for the most general possible harmonic oscillation of the gaseous layers as a whole. This system of equations (35)–(38) is the fundamental one upon which the subsequent arguments depend. A very important feature of this system is the difference of sign between the last sets of terms on the right-hand sides of the last two of the equations, suggesting the possibility of an accumulation of a difference in phase between the variations in displacement and temperature as we proceed through the model. It is pointed out in § 5, however, that these terms are only appreciable, in the case of stars of the masses, radii, temperatures, and periods of the Cepheids, in the layers lying close to the surface and so such a difference of phase is only appreciable in these layers. This is in agreement with Eddington's conclusions.‡ The value of this change of phase in these layers depends upon the values of certain constants of integration which have to be appropriately determined from the boundary conditions, and owing to the complexity of the system of equations can only be determined by trial through the very troublesome numerical integration of the system. It is significant to notice that when the extreme outside of the star is reached, the terms which were negligible in the interior become themselves the dominant terms in the equations (37) and (38), and so it is by no means certain that the phase change in the outer layers will be small, as Eddington, in the paper just quoted, has argued will be the case.

In § 6 we consider the form of the solution of (35), (36), (37), (38) very close to the boundary, and determine the number of arbitrary constants left in the solution when the physical conditions of continuity are inserted at the boundary. It is important to notice that

* loc. cit.

† Eddington, *M.N.* 77 (1916), 16, and later papers; Milne, *M.N.* 91 (1930), 4.

‡ *M.N.* 87 (1927), 539. See also J. J. M. Reesinck in *M.N.* 87 (1927), 414.

the boundary condition used by Eddington in his theory, namely, that $\delta P = 0$ at the surface, is automatically satisfied, since we show that this is a necessary consequence of the circumstance that the amplitude of the displacement at the boundary must be finite.

In § 7 we state the boundary conditions to be satisfied at the centre of the completely gaseous standard model configuration, this case providing the simplest illustration of the arguments adopted, although it has become of less significance in the consideration of actual stars occurring in Nature, since the advent of Milne's theory of stellar structure. It is established that these central boundary conditions are just sufficient to provide a determination of all the arbitrary constants left in the solution, together with a single equation which may be interpreted as determining the period, or periods, of the system in terms of the fundamental constants of the model. One of the conditions is, plainly, that there shall be no displacement at the centre. This is found to provide two equations between the constants left in the problem. It was suggested by Milne that a further condition to be satisfied at the centre is that there the star shall be hottest when most compressed, that is to say, the variations in displacement and temperature must be in phase at the centre. This was assumed tacitly in Eddington's theory and has not been explicitly stated previously as a condition of the required solution. This condition is fundamental in the present argument and provides the extra relation necessary for the determination of the period equation.

In a subsequent paper we hope to consider extensions of the method to more complicated examples.

2. The steady-state configuration. The equations of equilibrium of the standard model in which the opacity k and the rate of generation of energy per unit mass by internal processes ϵ are constant, are already well known but are restated here for reference.

Let M, L be the mass and luminosity, or total source strength, of the configuration respectively. Then if r denotes distance from the centre, p the gas pressure, p' the radiation pressure, ρ the density, T the temperature, $M(r)$ the mass within a sphere of radius r, $L(r)$ the total net flow of radiant energy across such a sphere per second,*

* The symbol $L(r)$ is used in two senses. It can be taken to mean the total flow of energy across the sphere per second, as defined here, or the total amount of energy being generated within the sphere per second. In a steady state these are, of course, equal, but we shall be dealing with non-steady states in

G the gravitational constant, and c the velocity of light, we have for a steady state

$$\frac{\partial p}{\partial r}+\frac{\partial p'}{\partial r}=-\frac{GM(r)}{r^2}\rho, \tag{1}$$

$$\frac{\partial p'}{\partial r}=-\frac{kL(r)}{4\pi cr^2}\rho, \tag{2}$$

and

$$\frac{L(r)}{M(r)}=\epsilon=\frac{L}{M}. \tag{3}$$

From these (assuming as boundary conditions $\rho=0$, $p'=0$) is derived the familiar relation

$$p=\frac{\beta}{1-\beta}p', \tag{4}$$

where

$$\beta=1-\frac{kL}{4\pi cGM}, \tag{5}$$

a constant of the model.

In the outer (gaseous) regions of the star the equation of state

$$p=\frac{R}{\mu}\rho T$$

holds, and from this and the relation

$$\frac{dM(r)}{dr}=4\pi r^2\rho$$

is derived the Emden 'polytropic' equation

$$\frac{1}{\xi^2}\frac{d}{d\xi}\left(\xi^2\frac{d\theta}{d\xi}\right)=-\theta^3, \tag{6}$$

in which the variables have been changed according to the relations

$$T=\lambda\theta, \tag{7}$$

and

$$r=\frac{1}{\lambda}\left(\frac{R}{\mu}\right)\frac{(1-\beta)^{\frac{1}{2}}}{\beta}\frac{1}{(\frac{1}{3}\pi aG)^{\frac{1}{2}}}\xi. \tag{8}$$

The 'discriminant' of the model is

$$\overline{C}=\frac{1-\beta}{\beta^4}\left(\frac{R}{\mu}\right)^4\frac{16}{\frac{1}{3}\pi aG^3M^2}, \tag{9}$$

a known function of L, M, and k.

which this is not necessarily true, and so it is necessary to keep the distinction clear and use the symbol only in the sense defined here.

If λ is so chosen that ξ_1, the value of ξ corresponding to $r = r_1$ the radius of the star, is equal to unity, then

$$\lambda = \left(\frac{R}{\mu}\right)\frac{(1-\beta)^{\frac{1}{2}}}{\beta}\,\frac{1}{(\frac{1}{3}\pi a G)^{\frac{1}{2}}}\,\frac{1}{r_1}, \tag{10}$$

and the boundary value of $d\theta/d\xi$ is given by

$$\left(\frac{d\theta}{d\xi}\right)_1 = -\frac{1}{C^{\frac{1}{2}}}. \tag{11}$$

If L has the critical value L_0 of Milne's theory the configuration is wholly gaseous and of the type investigated in Eddington's theory. Otherwise the configuration is either of the 'collapsed' type ($L < L_0$) or the 'centrally-condensed' type ($L > L_0$) of Milne's theory. We shall take note of this distinction later when considering the boundary conditions for the different models to be discussed. For the present we confine our attention to the derivation of the equations of motion of the gaseous layers of the star.

3. The equations of small motion of the gaseous layers. We fix attention on a particular element of material which in the equilibrium configuration is distant r_0 from the centre. The physical conditions of this element in the equilibrium configuration will be represented by symbols with zero suffixes attached. In the small motion to be considered we put

$$\left.\begin{aligned} r &= r_0+\delta r \\ T &= T_0+\delta T \\ \rho &= \rho_0+\delta\rho \end{aligned}\right\} \tag{12}$$

etc., where δr, δT, $\delta\rho$, etc., are small compared with r_0, T_0, ρ_0, etc., and are functions of r_0 and of the time t.

The equations to be satisfied throughout the small motion are:

(*a*) *a hydrodynamical equation of motion*

$$\frac{Du}{Dt} = -\frac{1}{\rho}\,\frac{\partial(p+p')}{\partial r} - \frac{GM(r)}{r^2}, \tag{13}$$

where u is the outward radial velocity and D/Dt denotes differentiation following the motion;

(*b*) *an equation of energy-exchanges*, expressing the fact that the rate of gain of energy per unit volume of any portion of the material, less the rate of loss of energy by radiation, is equal to the rate of gain of energy measured in terms of the physical properties of the

material: such an equation has been derived in a general form by Milne:* the form appropriate to radial motion of a spherical configuration is

$$\epsilon\rho-\frac{1}{r^2}\frac{\partial}{\partial r}\{r^2F(r)\}=\frac{DE}{Dt}-u\frac{\partial E}{\partial r}+c_v\rho\frac{DT}{Dt}-\frac{p+p'}{\rho}\frac{D\rho}{Dt}, \qquad (14)$$

where $F(r)$ is the net outward flux per unit area across a sphere of radius r, E the density of radiant energy per unit volume, and c_v the specific heat per unit volume of the material;

(c) *an equation of continuity.* This is obtained by expressing the fact that the mass within a small shell of radii r_0 and $r_0+\Delta r_0$ remains unaltered in the disturbed configuration, and is easily found to be

$$\frac{\delta\rho}{\rho_0}+\frac{2\delta r}{r_0}+\frac{\partial(\delta r)}{\partial r_0}=0. \qquad (15)$$

In substituting the expressions (12) into equations (13) and (14), the following lemma, which is capable of elementary proof, is of considerable use:

LEMMA. *If ϕ is any function of position in the sphere and $\phi=\phi_0$ when $r=r_0$, in the equilibrium configuration, and $\phi=\phi_0+\delta\phi$ when $r=r_0+\delta r$, in the disturbed configuration, then the space gradient of ϕ, $\partial\phi/\partial r$, in the disturbed configuration is equal to*

$$\left(\frac{\partial\phi}{\partial r}\right)_0\left(1-\frac{\partial(\delta r)}{\partial r_0}\right)+\frac{\partial(\delta\phi)}{\partial r_0},$$

neglecting quantities of the second and higher orders of smallness.

We can therefore write (13), to the first order of small quantities,

$$\frac{d^2(\delta r)}{dt^2}$$
$$=-\frac{1}{\rho_0}\left(1-\frac{\partial\rho}{\rho_0}\right)\left[\left(\frac{\partial(p+p')}{\partial r}\right)_0\left(1-\frac{\partial(\delta r)}{\partial r_0}\right)+\frac{\partial\delta(p+p')}{\partial r_0}\right]-\frac{GM(r)}{r_0^2}\left(1-\frac{2\delta r}{r_0}\right),$$

since $M(r)=[M(r)]_0$ always.

Using (1) and (15) this becomes

$$\frac{d^2(\delta r)}{dt^2}=-\frac{4}{\rho_0}\left(\frac{\partial(p+p')}{\partial r_0}\right)_0\frac{\delta r}{r_0}-\frac{1}{\rho_0}\frac{\partial\delta(p+p')}{\partial r_0}. \qquad (16)$$

* *Quart. J. of Math.* (Oxford), 1 (1930), 1. Other forms of this equation have been given by H. Vogt in *A.N.* 5545 (1928), 232, and L. H. Thomas in *Quart. J. of Math.* (Oxford), 1 (1930), 239.

Similarly, to the first order of small quantities, (14) can be written

$$(\epsilon_0+\delta\epsilon)(\rho_0+\delta\rho)-\frac{1}{4\pi r_0^2}\left(1-\frac{2\delta r}{r_0}\right)\left[\left(\frac{\partial L(r)}{\partial r}\right)_0\left(1-\frac{\partial(\delta r)}{\partial r_0}\right)+\frac{\partial\delta L(r)}{\partial r_0}\right] \tag{17}$$

$$=\frac{d(3p')}{dt}-3\left(\frac{\partial p'}{\partial r}\right)_0\frac{d(\delta r)}{dt}+\frac{c_v\rho_0 T_0}{4p_0'}\frac{d(\delta p')}{dt}-\frac{(p+p')_0}{\rho_0}\frac{d(\delta\rho)}{dt},$$

since $E=3p'$, $p'\propto T^4$, and $F(r)=L(r)/4\pi r^2$.

The coefficient of $(\partial L(r)/\partial r)_0$ on the left-hand side of (17) is, to the first order of small quantities,

$$1-\frac{2\delta r}{r_0}-\frac{\partial(\delta r)}{\partial r_0},$$

and this, by (15), is equal to

$$1+\frac{\delta\rho}{\rho_0}.$$

Also,

$$\left(\frac{\partial L(r)}{\partial r}\right)_0=\epsilon_0\left(\frac{\partial M(r)}{\partial r}\right)_0$$

$$=\epsilon_0 4\pi r_0^2\rho_0.$$

Hence the left-hand side of (17) reduces, to the first order, to

$$\epsilon_0\rho_0+\epsilon_0\delta\rho+\rho_0\delta\epsilon-\epsilon_0\rho_0\left(1+\frac{\delta\rho}{\rho_0}\right)-\frac{1}{4\pi r_0^2}\frac{\partial\delta L(r)}{\partial r_0},$$

that is, to

$$\rho_0\delta\epsilon-\frac{1}{4\pi r_0^2}\frac{\partial\delta L(r)}{\partial r_0}.$$

Further,

$$c_v\rho_0 T_0=\frac{p_0}{\gamma-1},$$

where

$$\gamma=c_p/c_v,$$

the specific heat ratio for the matter alone. Hence

$$\frac{c_v\rho_0 T_0}{4p_0'}=\frac{\beta}{4(\gamma-1)(1-\beta)},$$

from (4).

Hence we can now write (17) in the form

$$\frac{12(\gamma-1)(1-\beta)+\beta}{4(\gamma-1)(1-\beta)}\frac{d(\delta p')}{dt}$$

$$=3\left(\frac{\partial p'}{\partial r}\right)_0\frac{d(\delta r)}{dt}+\frac{(p+p')_0}{\rho_0}\frac{d(\delta\rho)}{dt}+\rho_0\delta\epsilon-\frac{1}{4\pi r_0^2}\frac{\partial(\delta L(r))}{\partial r_0}. \tag{18}$$

Owing to our lack of knowledge of the nature of the processes by which energy is liberated sub-atomically, we have no information as to the manner in which $\delta\epsilon$ will vary during the motion. We must here introduce an arbitrary hypothesis as to this variation. The

simplest suitable hypothesis to make is that $\delta\epsilon = 0$. The introduction of this hypothesis is equivalent to an averaging process for the energy liberation throughout the duration of the motion, necessitated by our lack of knowledge of the actual variation at each instant. It should be clearly understood that the hypothesis is quite distinct from that made in taking ϵ to be constant throughout the equilibrium configuration, which is equivalent to an averaging of the energy generation throughout the star. Each hypothesis represents an averaging process, the one in time, the other in space.

Our equations are then, finally,

$$\frac{d^2(\delta r)}{dt^2} = -\frac{4}{\rho}\frac{\partial(p+p')}{\partial r}\frac{\delta r}{r} - \frac{1}{\rho}\frac{\partial\delta(p+p')}{\partial r}, \tag{19}$$

$$\frac{12(\gamma-1)(1-\beta)+\beta}{4(\gamma-1)(1-\beta)}\frac{d(\delta p')}{dt} = \frac{3\partial p'}{\partial r}\frac{d(\delta r)}{dt} + \frac{p+p'}{\rho}\frac{d(\delta\rho)}{dt} - \frac{1}{4\pi r^2}\frac{\partial\delta L(r)}{\partial r}, \tag{20}$$

$$\frac{\delta\rho}{\rho} = -\left(\frac{2\delta r}{r} + \frac{\partial(\delta r)}{\partial r}\right), \tag{21}$$

in which we have dropped all the suffixes in the coefficients of the variational terms, it being understood henceforward that all such coefficients refer to the equilibrium values of the physical quantities concerned.

4. Conditions for harmonic oscillations of the gaseous layers. We seek the most general harmonic oscillation of the gaseous layers in which each element performs harmonic oscillations of period $2\pi/n$. Such an oscillation will be represented by the expressions

$$\left.\begin{aligned}\delta r &= \tilde{A}(r)\cos nt + \tilde{B}(r)\sin nt\\ \delta p' &= \tilde{C}(r)\cos nt + \tilde{D}(r)\sin nt\end{aligned}\right\}, \tag{22}$$

where the functions $\tilde{A}(r)$, $\tilde{B}(r)$ are small compared with r, and $\tilde{C}(r)$, $\tilde{D}(r)$ are small compared with p'. (A knowledge of these functions for the gaseous layers would determine the relative phases of the variations in displacement and net flux at any point; in particular, at the surface.)

Then, by (21),

$$\frac{\delta\rho}{\rho} = -\left\{\left(\frac{2\tilde{A}}{r} + \frac{d\tilde{A}}{dr}\right)\cos nt + \left(\frac{2\tilde{B}}{r} + \frac{d\tilde{B}}{dr}\right)\sin nt\right\}. \tag{23}$$

Also

$$\frac{\delta T}{T} = \frac{1}{4}\frac{\delta p'}{p'}$$

or
$$\frac{\delta T}{T} = \frac{1}{4p'}\{\tilde{C}\cos nt + \tilde{D}\sin nt\}. \tag{24}$$

Hence
$$\frac{\delta p}{p} = \frac{\delta\rho}{\rho} + \frac{\delta T}{T}$$
$$= \frac{\beta}{4(1-\beta)}\,\frac{\tilde{C}\cos nt + \tilde{D}\sin nt}{p} - \left\{\left(\frac{2\tilde{A}}{r} + \frac{d\tilde{A}}{dr}\right)\cos nt + \left(\frac{2\tilde{B}}{r} + \frac{d\tilde{B}}{dr}\right)\sin nt\right\},$$
using (4).

Hence
$$\delta(p+p') = \frac{4-3\beta}{4(1-\beta)}(\tilde{C}\cos nt + \tilde{D}\sin nt) -$$
$$-\frac{\beta}{1-\beta}p'\left\{\left(\frac{2\tilde{A}}{r} + \frac{d\tilde{A}}{dr}\right)\cos nt + \left(\frac{2\tilde{B}}{r} + \frac{d\tilde{B}}{dr}\right)\sin nt\right\}. \tag{25}$$

Also, from (2),
$$\frac{\delta L(r)}{L(r)} = \frac{2\delta r}{r} - \frac{\delta\rho}{\rho} + \delta\left(\frac{\partial p'}{\partial r}\right)\Big/\frac{\partial p'}{\partial r}$$
$$= \frac{4\delta r}{r} + \frac{\partial(\delta r)}{\partial r} + \left(\frac{\partial(\delta p')}{\partial r} - \frac{\partial p'}{\partial r}\,\frac{\partial(\delta r)}{\partial r}\right)\Big/\frac{\partial p'}{\partial r},$$
using (21) and the lemma of § 3. Hence
$$\frac{\delta L(r)}{L(r)} = \left(\frac{4\tilde{A}}{r} + \frac{d\tilde{C}}{dr}\Big/\frac{\partial p'}{\partial r}\right)\cos nt + \left(\frac{4\tilde{B}}{r} + \frac{d\tilde{B}}{dr}\Big/\frac{\partial p'}{\partial r}\right)\sin nt. \tag{26}$$

Substituting these expressions into equations (19) and (20), and equating the coefficients of $\cos nt$ and $\sin nt$ separately to zero, we find that the functions $\tilde{A}(r)$, $\tilde{B}(r)$, $\tilde{C}(r)$, $\tilde{D}(r)$ must satisfy the system of differential equations:
$$n^2\tilde{A} = \frac{4}{1-\beta}\,\frac{1}{\rho r}\,\frac{\partial p'}{\partial r}\,\tilde{A} + \frac{1}{\rho}\left[\frac{4-3\beta}{4(1-\beta)}\,\frac{d\tilde{C}}{dr} - \frac{\beta}{1-\beta}\,\frac{\partial}{\partial r}\left\{p'\left(\frac{2\tilde{A}}{r} + \frac{d\tilde{A}}{dr}\right)\right\}\right], \tag{27}$$
$$n^2\tilde{B} = \frac{4}{1-\beta}\,\frac{1}{\rho r}\,\frac{\partial p'}{\partial r}\,\tilde{B} + \frac{1}{\rho}\left[\frac{4-3\beta}{4(1-\beta)}\,\frac{d\tilde{B}}{dr} - \frac{\beta}{1-\beta}\,\frac{\partial}{\partial r}\left\{p'\left(\frac{2\tilde{B}}{r} + \frac{d\tilde{B}}{dr}\right)\right\}\right], \tag{28}$$
$$\frac{12(\gamma-1)(1-\beta)+\beta}{4(\gamma-1)(1-\beta)}n\tilde{C}$$
$$= 3n\,\frac{\partial p'}{\partial r}\,\tilde{A} - \frac{np'}{1-\beta}\left(\frac{2\tilde{A}}{r} + \frac{d\tilde{A}}{dr}\right) - \frac{c}{kr^2}\,\frac{\partial}{\partial r}\left\{\frac{r^2}{\rho}\left(\frac{4\tilde{B}}{r}\,\frac{\partial p'}{\partial r} + \frac{d\tilde{D}}{dr}\right)\right\}, \tag{29}$$
$$\frac{12(\gamma-1)(1-\beta)+\beta}{4(\gamma-1)(1-\beta)}n\tilde{D}$$
$$= 3n\,\frac{\partial p'}{\partial r}\,\tilde{B} - \frac{np'}{1-\beta}\left(\frac{2\tilde{B}}{r} + \frac{d\tilde{B}}{dr}\right) + \frac{c}{kr^2}\,\frac{\partial}{\partial r}\left\{\frac{r^2}{\rho}\left(\frac{4\tilde{A}}{r}\,\frac{\partial p'}{\partial r} + \frac{d\tilde{C}}{dr}\right)\right\}. \tag{30}$$

We express these equations in terms of the Emden variables defined in § 2. We write

$$\tilde{A}(r) = \frac{1}{\lambda}\left(\frac{R}{\mu}\right)\frac{(1-\beta)^{\frac{1}{2}}}{\beta}\frac{1}{(\frac{1}{3}\pi a G)^{\frac{1}{2}}}A(\xi),$$

or, by (10),

$$\tilde{A}(r) = r_1 A(\xi), \tag{31}$$

where r_1 is the radius of the star, and, similarly,

$$\tilde{B}(r) = r_1 B(\xi). \tag{31'}$$

Also, put

$$\tilde{C}(r) = \tfrac{1}{3}a\lambda^4 C(\xi), \tag{32}$$

and

$$\tilde{D}(r) = \tfrac{1}{3}a\lambda^4 D(\xi). \tag{32'}$$

We further introduce new symbols ω^2 and Δ defined by

$$\omega^2 = \frac{1}{4\lambda}\frac{1}{(R/\mu)}\frac{\beta}{1-\beta}r_1^2,$$

or, in terms of the discriminant of the equilibrium model,

$$\omega^2 = \frac{1}{1-\beta}\frac{r_1^3}{GM\bar{C}^{\frac{1}{2}}}; \tag{33}$$

and

$$\Delta = \frac{c}{k}\frac{1}{r_1^2}\frac{1-\beta}{\beta}\frac{(R/\mu)}{\frac{1}{3}a\lambda^3},$$

or

$$\Delta = \frac{1}{1-\beta}\frac{Lr_1}{GM^2\bar{C}^{\frac{1}{2}}}. \tag{34}$$

(We observe that $n^2\omega^2$ and Δ/n are dimensionless numbers.)

Then, using the relations

$$p' = \tfrac{1}{3}a\lambda^4\theta^4$$

and

$$\rho = \frac{1}{3}\frac{a\lambda^3}{(R/\mu)}\frac{\beta}{1-\beta}\theta^3,$$

the equations (27), (28), (29), (30) become (dashes denoting differentiations with respect to ξ)

$$\frac{4-3\beta}{4(1-\beta)}C' = \left(\omega^2 n^2 - \frac{4}{1-\beta}\frac{\theta'}{\xi}\right)\theta^3 A + \frac{\beta}{4(1-\beta)}\frac{\partial}{\partial\xi}\left\{\theta^4\left(\frac{2A}{\xi}+A'\right)\right\}, \tag{35}$$

$$\frac{4-3\beta}{4(1-\beta)}D' = \left(\omega^2 n^2 - \frac{4}{1-\beta}\frac{\theta'}{\xi}\right)\theta^3 B + \frac{\beta}{4(1-\beta)}\frac{\partial}{\partial\xi}\left\{\theta^4\left(\frac{2B}{\xi}+B'\right)\right\}, \tag{36}$$

$$\frac{12(\gamma-1)(1-\beta)+\beta}{4(\gamma-1)(1-\beta)}C = 3\theta^3\theta' A - \frac{\theta^4}{4(1-\beta)}\left(\frac{2A}{\xi}+A'\right) - \frac{\Delta}{n}\frac{1}{\xi^2}\frac{\partial}{\partial\xi}\left\{\xi^2\left(\frac{4\theta' B}{\xi}+\frac{D'}{\theta^3}\right)\right\}, \tag{37}$$

$$\frac{12(\gamma-1)(1-\beta)+\beta}{4(\gamma-1)(1-\beta)}D = 3\theta^3\theta' B - \frac{\theta^4}{4(1-\beta)}\left(\frac{2B}{\xi}+B'\right)+\frac{\Delta}{n}\frac{1}{\xi^2}\frac{\partial}{\partial\xi}\left\{\xi^2\left(\frac{4\theta' A}{\xi}+\frac{C'}{\theta^3}\right)\right\}. \quad (38)$$

The variations in the various physical quantities involved are found to be, in virtue of equations (22), (23), (25), (26), and the relations between the new and old variables,

$$\delta r = r_1(A\cos nt + B\sin nt), \quad (39)$$

$$\delta p' = \frac{1-\beta}{4\pi}\frac{GM^2\bar{C}}{r_1^4}(C\cos nt + D\sin nt), \quad (40)$$

$$\delta\rho = -\frac{M\bar{C}^{\frac{1}{2}}\theta^3}{4\pi r_1^3}\left\{\left(\frac{2A}{\xi}+A'\right)\cos nt+\left(\frac{2B}{\xi}+B'\right)\sin nt\right\}, \quad (41)$$

$$\delta(p+p') = \frac{GM^2\bar{C}}{16\pi r_1^4}\Big[(4-3\beta)(C\cos nt + D\sin nt) - \beta\theta^4\left\{\left(\frac{2A}{\xi}+A'\right)\cos nt+\left(\frac{2B}{\xi}+B'\right)\sin nt\right\}\Big] \quad (42)$$

$$\delta L(r) = -\frac{L\bar{C}^{\frac{1}{2}}}{4\pi}\left\{\left(\frac{4\theta' A}{\xi}+\frac{C'}{\theta^3}\right)\cos nt+\left(\frac{4\theta' B}{\xi}+\frac{D'}{\theta^3}\right)\sin nt\right\}. \quad (43)$$

The equations (35), (36), (37), (38) are four linear differential equations for the functions A, B, C, D. The solution of the problem of the small oscillations of the gaseous layers of a standard model configuration consists in the determination of a solution of these equations, which satisfies appropriate boundary conditions according to the model under consideration. These conditions for certain particular models are considered in later paragraphs. Such a solution being determined, the physical changes of the configuration during the motion are completely determined by equations (39), (40), (41), (42), (43).

5. Simplification of the equations in the interior of the star. The last sets of terms on the right-hand sides of (37) and (38) have, it should be noticed, opposite signs. These terms arise from the term involving $\delta L(r)$ in (30), and hence will be for convenience referred to as the 'flow' terms. The difference of sign in the flow terms in equations (37) and (38) suggests the possibility of an accumulation of phase difference between the variations in displacement and luminosity as we proceed through the star. But it must be

noticed that for a star whose internal temperature is of the order of several million degrees, oscillating with a period of the order of a few days, the quantity Δ/n, defined by (34), is very small. So long as θ^3 is large in comparison with Δ/n, which is true for a very great part of the star, and fails only for a very small fringe at the surface, the flow terms in (37) and (38) are negligible. When we neglect these terms, the system of equations falls naturally into two pairs, one involving only the functions A and C, and the other involving only B and D, and the latter pair is the same as the former except for the replacement of A by B and of C by D. In these circumstances $A/C = B/D$ for all values of ξ; that is to say, there is no change of phase between δr and $\delta p'$ in these regions of the star. Any such change is due to the effect of the flow terms in (37) and (38) and will therefore occur only in the outer fringe of the star when these flow terms become significant. This is in agreement with the results of Eddington's theory of 'adiabatic' pulsations.

6. Form of the solution of the equations near the surface. The system of equations (35), (36), (37), (38) consists of four linear homogeneous equations, each of the second order. The most general solution will therefore contain eight arbitrary constants of integration. Eight 'boundary' conditions will be required to determine these, and one further will be necessary to provide a period equation for the system; this latter equation is, then, the eliminant of the eight constants from nine equations. It is not at once obvious, however, how many conditions are represented by the circumstance that the amplitude of the oscillation must remain finite at the boundary, i.e. that the functions A and B in (35), (36), (37), (38) must be continuous at the boundary. It is therefore first necessary to investigate the most general continuous form of the solution of these equations very close to the boundary, $\xi = 1$, with a view to discovering the degree of indeterminacy left in the solution when the condition of continuity of the amplitude at the surface has been satisfied.

Let $\eta = 1-\xi$, so that η is small near the surface.

Then it can easily be shown that the form of θ, the solution of (6), near the surface, is

$$\theta = \frac{1}{C^{\frac{1}{4}}}\{\eta+\eta^2+\eta^3+\eta^4+O(\eta^5)\}.$$

Then, if A and B are continuous at the surface, the right-hand

sides of equations (35) and (36) are at least $O(\eta^3)$ very close to the surface.

Hence C' and D' are at least $O(\eta^3)$ very close to the surface; hence C and D are at least $O(\eta^4)$ very close to the surface.

(It follows from (42) that, at the boundary, δP is of necessity zero, if A and B are continuous there. The condition $\delta P = 0$ cannot therefore be used, as Eddington uses it, as a boundary condition to determine the appropriate value of n^2 for a physically suitable solution, since it is a logical consequence of the equations of motion themselves, together with the physical property of continuity of the solution at the boundary.)

The left-hand sides of (37) and (38) and the right-hand sides with the exception of the flow terms are all therefore $O(\eta^3)$ at least. Hence, sufficiently close to the boundary, these flow terms become dominant, however small Δ/n may be. ($\Delta/n = 0$ implies n infinite.)

Hence, sufficiently close to the boundary, we can neglect all but the flow terms in (37) and (38), so that

$$\frac{\partial}{\partial\xi}\left\{\xi^2\left(\frac{4\theta' B}{\xi}+\frac{D'}{\theta^3}\right)\right\} = 0,$$

and

$$\frac{\partial}{\partial\xi}\left\{\xi^2\left(\frac{4\theta' A}{\xi}+\frac{C'}{\theta^3}\right)\right\} = 0.$$

Hence

$$\frac{4\theta' B}{\xi}+\frac{D'}{\theta^3} = \frac{k_B}{\xi^2},$$

and

$$\frac{4\theta' A}{\xi}+\frac{C'}{\theta^3} = \frac{k_A}{\xi^2},$$

where k_A and k_B are arbitrary constants.

Further, the dominant terms in (35), (36), namely, those in θ^3, give

$$\frac{4-3\beta}{4(1-\beta)}\frac{C'}{\theta^3} = \left(\omega^2 n^2 - \frac{4}{1-\beta}\frac{\theta'}{\xi}\right)A + \frac{\beta}{1-\beta}\theta'\left(\frac{2A}{\xi}+A'\right),$$

and

$$\frac{4-3\beta}{4(1-\beta)}\frac{D'}{\theta^3} = \left(\omega^2 n^2 - \frac{4}{1-\beta}\frac{\theta'}{\xi}\right)B + \frac{\beta}{1-\beta}\theta'\left(\frac{2B}{\xi}+B'\right).$$

Hence, sufficiently close to the boundary of the star, the function A must satisfy the equation

$$\frac{4-3\beta}{4(1-\beta)}\frac{4\theta' A}{\xi} + \left(n^2\omega^2 - \frac{4}{1-\beta}\frac{\theta'}{\xi}\right)A + \frac{\beta}{1-\beta}\theta'\left(\frac{2A}{\xi}+A'\right) = \frac{4-3\beta}{4(1-\beta)}\frac{k_A}{\xi^2},$$

and the same is true of B, k_B replacing k_A. Hence,

$$\theta' A' + \left(\frac{1-\beta}{\beta} n^2\omega^2 - \frac{\theta'}{\xi}\right) A = \frac{4-3\beta}{4\beta} \frac{k_A}{\xi^2}, \tag{44}$$

and

$$\theta' B' + \left(\frac{1-\beta}{\beta} n^2\omega^2 - \frac{\theta'}{\xi}\right) B = \frac{4-3\beta}{4\beta} \frac{k_B}{\xi^2}. \tag{45}$$

In the region under consideration the form of θ, given by the solution of (6), is

$$\theta = \frac{1}{\bar{C}^{\frac{1}{2}}}(\eta + \eta^2 + \eta^3 + \eta^4 + O(\eta^5)),$$

where

$$\xi = 1 - \eta.$$

so that η is very small indeed. Thus, to a close approximation,

$$\theta = \frac{1}{\bar{C}^{\frac{1}{2}}} \frac{\eta}{1-\eta},$$

or

$$\theta = \frac{1}{\bar{C}^{\frac{1}{2}}} \frac{1-\xi}{\xi}.$$

Then

$$\theta' = -\frac{1}{\bar{C}^{\frac{1}{2}}} \frac{1}{\xi^2}.$$

Hence we may rewrite (44) and (45) in the form

$$A' - \left(\frac{1-\beta}{\beta} n^2\omega^2 \bar{C}^{\frac{1}{2}} \xi^2 + \frac{1}{\xi}\right) A = -\bar{C}^{\frac{1}{2}} \frac{4-3\beta}{4\beta} k_A \tag{44'}$$

$$B' - \left(\frac{1-\beta}{\beta} n^2\omega^2 \bar{C}^{\frac{1}{2}} \xi^2 + \frac{1}{\xi}\right) B = -\bar{C}^{\frac{1}{2}} \frac{4-3\beta}{4\beta} k_B. \tag{45'}$$

Integrating these equations we find that the most general continuous form of the functions A and B very close to the surface of the star can be written

$$A = \xi^{-1} e^{-\sigma\xi^3} \left\{ K_A - \bar{C}^{\frac{1}{2}} \frac{4-3\beta}{4\beta} k_A \int \xi e^{\sigma\xi^3} \, d\xi \right\},$$

$$B = \xi^{-1} e^{-\sigma\xi^3} \left\{ K_B - \bar{C}^{\frac{1}{2}} \frac{4-3\beta}{4\beta} k_B \int \xi e^{\sigma\xi^3} \, d\xi \right\},$$

where K_A and K_B are undetermined constants and

$$\sigma = \frac{1-\beta}{\beta} n^2\omega^2 \bar{C}^{\frac{1}{2}}.$$

The solution thus contains *four* arbitrary constants, k_A, k_B, K_A,

K_B. We have as yet made no stipulation concerning our origin of time. If we so choose this origin as to fix the variation in energy-output at the surface in value and phase, we at once assign values to k_A and k_B. Thus, if we choose the origin of time so that the variation of $L(r)$ at the surface, i.e. the *observed* variation in brightness, is a pure $\cos nt$ variation, we can put $k_A = 0$ and assign a definite value to k_B according to the amplitude of the variation in brightness. Suppose, then, we choose $k_A = 0$ and $k_B = 1$. Then the solution contains *two* remaining undetermined constants, K_A and K_B. Knowledge of the ratio of these two constants would give us the phase of the variation in displacement at the boundary. For when $\xi = 1$,

$$A = e^{-\sigma} K_A,$$

and

$$B = e^{-\sigma} K_B,$$

since we may, without loss of generality, assume the additive constants for the indefinite integrals to be such as to make these integrals vanish for $\xi = 1$. But the ratio cannot be determined by consideration of the surface layers alone. The constants K_A and K_B can only be evaluated by finding which solution of (44) and (45) can be fitted on to the solutions of the equation in the interior regions where the present approximations break down.

We have now considered the solutions of the general system of differential equations for the motion of the sphere in two different regions—in the interior regions where the flow terms are negligible, and in the surface layers where the flow terms are dominant. The former regions extend throughout almost the whole sphere; the latter occupy but a small fringe at the boundary. Between these two regions there lies a transition region where the flow terms are comparable in magnitude with the other terms and where the equations of motion do not admit of simplification. The intractable nature of the equations in this region leaves a gap which cannot be bridged mathematically and makes it difficult to link up the appropriate solution for the interior with that for the surface. We are thus unable to determine quantitatively the phase relation between the variations in displacement and temperature at the surface. As we have already remarked, these variations are in phase in the interior or 'adiabatic' region. It is in the transitional region that a phase difference begins to accumulate, but by how much we cannot say. K_A and K_B can only be determined by continuing the solution for the interior

through the transitional layer to the surface region, and though this is possible theoretically, the nature of the equations makes it impracticable.

The region in which the approximations (44) and (45) are legitimate depends upon Δ/n. The necessary criterion is that θ^3 shall be negligible in comparison with Δ/n. Now for spheres of the masses, radii, and periods of the Cepheids, the value of Δ/n is about 10^{-6}. Hence θ must be negligible in comparison with 10^{-2}. Thus, for example, to four significant figures the approximations are legitimate only if $\theta < 10^{-6}$, and this is true for so minute a fringe at the edge of the star that the approximations described above are of no value in the actual solution of the equations, although, as we have seen, they indicate the degree of indeterminacy left in the problem when the surface conditions have been chosen appropriately.

7. Internal boundary conditions for a gaseous sphere. As already indicated, the determination of K_A and K_B requires the choice of an appropriate solution in the interior region. It therefore remains to consider what internal boundary conditions must be satisfied for certain models. In the present paper we confine our attention to the simplest model available—the completely gaseous sphere (Eddington model).* In Part II of this paper more complicated models, based on the recent researches of Milne, will be considered.

In the present case the conditions to be considered are those at the centre, $\xi = 0$.

In the first place it is obvious that two conditions to be fulfilled there are

$$A(0) = 0, \tag{46}$$

and

$$B(0) = 0, \tag{47}$$

since the centre must remain stationary.

A further condition that must be satisfied at the centre is that the phases of the variations in displacement and temperature as the centre is approached will be such that the instant of greatest compression there will coincide with that of highest temperature.† There is, of course, no direct evidence for this, but physical intuition is

* Jeans has claimed in the series of papers quoted in the Introduction to this paper that this model is vibrationally unstable.

† Suggested to me by Professor Milne.

strongly in favour of its validity in the case of matter at rest, generating energy at a constant rate (since we have assumed that $\delta\epsilon = 0$), while subject to periodic variations of compression.

Now, $$\delta r \propto A\cos nt + B\sin nt,$$

and $$\delta T \propto C\cos nt + D\sin nt.$$

Hence $$\frac{d(\delta r)}{dt} \propto n(-A\sin nt + B\cos nt),$$

and $$\frac{d(\delta T)}{dt} \propto n(-C\sin nt + D\cos nt).$$

Hence, greatest and least compressions occur at the centre when

$$\tan nt = B/A.$$

Similarly, greatest and least temperatures occur when

$$\tan nt = D/C.$$

Hence the above condition requires that

$$\lim_{\phi\to 0}\frac{B}{A} = \left(\frac{D}{C}\right)_{\xi=0},$$

or $$\frac{B'}{A'} = \frac{D}{C} \text{ at } \xi = 0. \tag{48}$$

The equations (46), (47), (48) constitute a set of three relations of the type $F_i(n, \beta, \gamma, \overline{C}^{\frac{1}{4}}, K_A K_B) = 0 \qquad (i = 1, 2, 3).$

From them K_A, K_B can be eliminated, leaving a relation of the type

$$\phi(n, \beta, \gamma, \overline{C}^{\frac{1}{4}}) = 0. \tag{49}$$

where ϕ represents a function not necessarily single-valued.

This is the period equation whose existence we desired to establish, appropriate solutions for the system being obtained only for such values of n as satisfy this relation. If this equation has its roots in n all real, the only possible motions are stable ones, but if imaginary roots in n are possible, unstable motions may occur and the model is unstable vibrationally. As (49) cannot be written in explicit form from the present analysis, no examination of the nature of its roots is here possible.

Further, when an appropriate value of n has been chosen according to (49), any two of the equations (46), (47), (48) determine corresponding values of K_A and K_B. This determines the phase relation

between the variations in displacement and temperature (and therefore also in luminosity, from (43)) at the boundary. It is therefore established that a definite phase relation does exist at the boundary between these variations. As stated earlier, this could only be determined through the long and excessively complicated numerical quadrature of the fundamental system of equations for specimen values of the constants involved (β and γ).

Note. In the course of preparation of the present paper for publication there has appeared a thorough investigation by S. Rosseland* of the nature of the stability criteria for gaseous spheres subject to radial pulsations. The procedure employed is essentially the same as that adopted here. (The equation of energy-change used by Rosseland differs from Milne's equation, being identical with that used by both Jeans and Vogt in their earlier investigations, quoted at the commencement of this paper.) While Rosseland's results are, in the main, directed towards stability considerations, and, for stars built on the standard model, corroborate the conclusions of Jeans, obtained by less elegant methods, his investigations of the atmospheric problem lead to results essentially the same as those of the present paper as regards the extent of the 'adiabatic' region of the star and the accumulation of phase difference between the variations in displacement and temperature in the transitional region just below the surface when these equations of motion do not admit of simplification.

The development of Rosseland's paper follows the lines of the general dynamical theory of oscillating systems and goes far towards supplying the need, mentioned in the Introduction to this paper, of a rational theory of stellar oscillations. Thus, for example, the orthogonality property of the oscillational functions for the amplitudes is derived for certain systems. Also an estimate is made of the order of magnitude of the period of the fundamental mode of oscillation of the star. The establishment of a period relation by the detailed consideration of boundary conditions, which is the main feature of the present paper, does not, however, form a part of Rosseland's investigation.

* 'On the stability of gaseous stars': *Univ. Obs. Oslo Pub.* No. 1 (1931).

A RELATION BETWEEN HYPERGEOMETRIC SERIES

By J. L. BURCHNALL (*Durham*)

[Received 9 November 1932]

MR. H. B. C. DARLING has recently* given the analogues, for functions of higher order, of the formula

$$F\left(\begin{matrix}a, b;\\ c\end{matrix}x\right)F\left(\begin{matrix}1-a, 1-b;\\ 2-c\end{matrix}x\right) = F\left(\begin{matrix}a+1-c, b+1-c;\\ 2-c\end{matrix}x\right)F\left(\begin{matrix}c-a, c-b;\\ c\end{matrix}x\right).$$

The machinery he employs increases in complexity as the order of the functions increases and does not readily establish the general result suggested by the particular instances. I venture therefore to give a general proof from which it will appear that the theory of adjoint operators indicates the existence of relations of the type in question and that a simple partial-fraction theorem suffices to determine their precise form.

In the interests of symmetry I slightly modify the conventional notation. Let $c_1, c_2, \ldots, c_n$ be numbers incongruent to modulus unity of which one, say c_r, is unity. Then I write

$$F\left(\begin{matrix}a_1, a_2, \ldots, a_n;\\ c_1, c_2, \ldots, c_{r-1}, c_{r+1}, \ldots, c_n\end{matrix}x\right) \equiv G\left(\begin{matrix}a_1, a_2, \ldots, a_n;\\ c_1, c_2, \ldots, c_n\end{matrix}x\right) \equiv G(a, c, x). \quad (1)$$

Let δ denote xd/dx, and notice that the operator adjoint to $\delta+a$ is $\delta+1-a$.

Consider the equation

$$\prod_{r=1}^{n}(\delta+b_r-1)y = x\prod_{r=1}^{n}(\delta+a_r)y. \quad (2)$$

Its adjoint is $\prod_{r=1}^{n}(\delta+2-b_r)w = \prod_{r=1}^{n}(\delta+1-a_r)xw.$

If $z = xw$, then z satisfies

$$\prod_{r=1}^{n}(\delta+2-b_r-1)z = x\prod_{r=1}^{n}(\delta+1-a_r)z. \quad (3)$$

Linearly independent solutions of (2), (3) are, for $r = 1, 2, \ldots, n$,

$$\left.\begin{aligned} y_r &= x^{1-b_r}G(a+1-b_r, b+1-b_r, x) \equiv x^{1-b_r}\sum_{m=0}^{\infty}A_{rm}x^m \\ z_r &= x^{b_r-1}G(b_r-a, b_r-b+1, x) \equiv x^{b_r-1}\sum_{m=0}^{\infty}B_{rm}x^m \end{aligned}\right\}, \quad (4)$$

* *Proc. London Math. Soc.* (2), 34 (1932), 323.

where

$$\left.\begin{aligned} A_{rm} &= \prod_{s=1}^{n} [(a_s+1-b_r)_m/(b_s+1-b_r)_m] \\ B_{rm} &= \prod_{s=1}^{n} [(b_r-a_s)_m/(b_r-b_s+1)_m] \end{aligned}\right\}. \tag{5}$$

Now, by the general theory of the adjoint,* there exists an identity of the form

$$\sum_{r,s} \lambda_{rs} y_r z_s \equiv 0,$$

and consideration of the various terms in the neighbourhood of the origin makes it clear that the identity is of the form

$$\sum_{r} \mu_r y_r z_r \equiv 0. \tag{6}$$

More precisely I show that, if $f(t) \equiv \prod_{r=1}^{n} (t-b_r)$, then

$$\sum_{r=1}^{n} [y_r z_r / f'(b_r)] \equiv 0. \tag{7}$$

It is evident that (7) is established if, for every k,

$$\sum_{r=1}^{n} \Big[\Big(\sum_{m=0}^{k} A_{r,k-m} B_{rm} \Big) \Big/ f'(b_r) \Big] = 0. \tag{8}$$

To obtain (8) write $\quad g(t) \equiv \prod_{r=1}^{n} (t-a_r)$

and consider the 'key-function'

$$H(t) \equiv \frac{g(t-1)g(t-2)\dots g(t-k)}{f(t)f(t-1)\dots f(t-k)}. \tag{9}$$

The analysis has proceeded on the assumption that the b_r are incongruent to modulus unity, so the denominator of $H(t)$ has no repeated roots. Moreover, the degree of the denominator exceeds that of the numerator by at least two, since n is assumed greater than unity. Thus the sum of the residues of $H(t)$ at its various poles is zero.

Now the residue at $t = b_r+m$ $(m < k)$ is

$$\frac{\prod_{l=0}^{m-1} \Big[\prod_{s=1}^{n} (b_r+l-a_s) \Big] \prod_{l=1}^{k-m} \Big[\prod_{s=1}^{n} (b_r-l-a_s) \Big]}{\prod_{l=1}^{m} \Big[\prod_{s=1}^{n} (b_r+l-b_s) \Big] f'(b_r) \prod_{l=1}^{k-m} \Big[\prod_{s=1}^{n} (b_r-l-b_s) \Big]}$$

or

$$\prod_{s=1}^{n} \left[\frac{(b_r-a_s)_m (a_s+1-b_r)_{k-m}}{(b_r-b_s+1)_m (b_s+1-b_r)_{k-m}} \right] \frac{1}{f'(b_r)},$$

which is $(A_{r,k-m} B_{r,m})/f'(b_r)$, and, on summing over all the poles, (8) is established.

* See, e.g., Darboux, *Théorie générale des surfaces*, ii, p. 103.

It is to be noted that the theory of the adjoint* requires also the following relations between the y_r and the ith derivatives $z_r^{(i)}$ of the z_r, namely,

$$\left.\begin{aligned}\sum_{r=1}^{n}[y_r z_r^{(i)}/f'(b_r)] &= 0 \qquad (i<n-1)\\ &= Cx^{1-n}(1-x)^{-1} \qquad (i=n-1)\end{aligned}\right\}, \tag{10}$$

where C is a numerical constant. On picking out the coefficient of x^{1-n} on the left of (10), we have

$$C=\sum_{r=1}^{n}\frac{(b_r-1)(b_r-2)\ldots(b_r-n+1)}{f'(b_r)}=1, \tag{11}$$

by a well-known theorem in partial fractions.

The 'key-function' corresponding to (10) is

$$H_i(t)\equiv(t-1)(t-2)\ldots(t-i)H(t).$$

Since the derivative of a hypergeometric series is a multiple of a hypergeometric series, the relations (10) lead to a great variety of relations between such series. The simplest is, in the usual notation,

$$\begin{aligned}&\frac{(1-a)(1-b)}{(1-c)(2-c)}F\left(\begin{matrix}a,b;\\ c\end{matrix}x\right)F\left(\begin{matrix}2-a,2-b;\\ 3-c\end{matrix}x\right)+\\ &+\frac{(c-a)(c-b)}{c(c-1)}F\left(\begin{matrix}a+1-c,b+1-c;\\ 2-c\end{matrix}x\right)F\left(\begin{matrix}c-a+1,c-b+1;\\ c+1\end{matrix}x\right)+\\ &+x^{-1}F\left(\begin{matrix}a+1-c,b+1-c,\\ 2-c\end{matrix}x\right)F\left(\begin{matrix}c-a,c-b;\\ c\end{matrix}x\right)\\ &\qquad\qquad=x^{-1}(1-x)^{-1}.\end{aligned} \tag{12}$$

* Darboux, loc. cit.

HEFFER'S FOR ALL BOOKS

Scientific Books and Journals of all kinds, New and Second-hand, English and Foreign. Heffer's are always prepared to purchase complete sets or long runs of the *Quarterly Journal of Mathematics*. Catalogue 386, mainly Mathematics, Physics, Chemistry, and Publications of Learned Societies, free on request.

W. HEFFER & SONS, LTD., CAMBRIDGE, ENGLAND

OXFORD BOOKS

WAVE MECHANICS
The Elementary Theory
By J. FRENKEL
20/- net

THE THEORY OF ELECTRIC AND MAGNETIC SUSCEPTIBILITIES
By J. H. VAN VLECK
30/- net

". . . There is throughout a welcome absence of dogmatism and of that suggestion of finality which is so often irritatingly conveyed in theoretical works; moreover, excellent reviews are given of experimental results. The book is a most valuable contribution to the study of magnetic and electric susceptibilities, and to Quantum-Mechanical Theory. . . ."—*Nature*

THE PRINCIPLES OF QUANTUM MECHANICS
By P. A. M. DIRAC
17/6 net

". . . There can be no doubt that his work ranks as one of the high achievements of contemporary physics. . . ."—*Nature*

THE CORRESPONDENCE AND PAPERS OF EDMOND HALLEY
By E. F. MacPIKE
21/- net

". . . This beautifully printed volume gives an admirable glimpse into the scientific activities of one of the most famous of English astronomers. . . ."—*Cambridge Review*

OXFORD UNIVERSITY PRESS

DEIGHTON, BELL & CO., LTD.

13 TRINITY STREET, CAMBRIDGE

UNIVERSITY BOOKSELLERS

ENGLISH AND FOREIGN

We supply

NEW AND SECOND-HAND TEXT-BOOKS in all Subjects

FOREIGN BOOKS in all Languages

The LATEST PUBLICATIONS in all Branches of Literature

OLD AND RARE BOOKS, &c.

We buy

LIBRARIES or SMALLER COLLECTIONS OF SCIENTIFIC BOOKS

JOURNALS, PUBLICATIONS OF LEARNED SOCIETIES, &c.

English and Foreign

Orders by post receive prompt and expert attention
Catalogues issued regularly

Telephone: CAMBRIDGE 239 ESTABLISHED 1700

Published in 1932

FOUNDATIONS OF POINT SET THEORY. By R. L. MOORE. American Mathematical Society Colloquium Publications. Volume XIII. 7 + 486 pp. $5.00

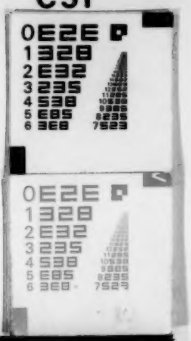
DIFFERENTIAL EQUATIONS FROM THE ALGEBRAIC STANDPOINT. By J. F. RITT. American Mathematical Society Colloquium Publications. Volume XIV. 10 + 172 pp. $2.50

All volumes of the Colloquium Series may be ordered from the American Mathematical Society, 501 West 116th Street, New York, or from Bowes & Bowes, 1 Trinity Street, Cambridge, England.

Printed in Great Britain at the University Press, Oxford, by John Johnson, Printer to the University

JL

C51


0E2E
1328
2E32
3235
4538
5E85
6 3E8
7523
0E2E
1328
2E32
3235
4538
5E85
6 3E8-
7523